

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 09:25:55 ; Search time 1999 Seconds

(without alignments)  
7110.754 Million cell updates/sec

Title: US-10-020-540A-1

Perfect score: 476

Sequence: 1 aactatgctgatgacaagat.....aacctccctgcttacaacc 476

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size: 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	4.6	279	29	CE765480 tigr-gss-
2	21	4.4	373	12	BG359460 sac24e07.
3	21	4.4	594	28	AZ756615 evl3h12.r
4	21	4.4	595	28	AZ756578 evl3e08.r

5	21	4.4	602	28	AZ756590
6	21	4.4	772	12	BG483395
7	21	4.4	826	29	CC969844
8	20	4.2	196	10	BE949260
9	20	4.2	216	9	AA759629
10	20	4.2	247	10	AW913044
11	20	4.2	287	14	T38551
12	20	4.2	294	10	BF399080
13	20	4.2	317	10	BB747879
14	20	4.2	318	10	BF463976
15	20	4.2	331	9	AA410171
16	20	4.2	362	12	EG792711
17	20	4.2	362	12	EG817350
18	20	4.2	372	10	BB791337
19	20	4.2	385	9	AA250251
20	20	4.2	390	9	AA083503
21	20	4.2	397	13	BY638039
22	20	4.2	399	13	BY404218
23	20	4.2	400	13	BY656976
24	20	4.2	401	13	C87655
25	20	4.2	405	13	BY631080
26	20	4.2	408	13	BY646779
27	20	4.2	408	13	BY656917
28	20	4.2	409	9	AA396378
29	20	4.2	415	13	BY612244
30	20	4.2	417	10	BF147358
31	20	4.2	419	13	BY623467
32	20	4.2	420	10	BB786872
33	20	4.2	421	12	BM250201
34	20	4.2	423	9	AA112100
35	20	4.2	424	10	BB781226
36	20	4.2	424	13	BY640680
37	20	4.2	425	10	BB743084
38	20	4.2	427	9	AA983024
39	20	4.2	429	13	BY535622
40	20	4.2	442	10	BB782299
41	20	4.2	445	14	CA885627
42	20	4.2	450	10	BB796119
43	20	4.2	455	9	AA959669
44	20	4.2	457	13	BY638518
45	20	4.2	458	10	BB795665

#### ALIGNMENTS

RESULT 1  
CE765480  
LOCUS  
DEFINITION  
tigr-gss-dog-17000370758071 Dog Library Canis familiaris genomic,  
genomic survey sequence.  
ACCESSION  
CE765480.1 GI:37106244  
VERSION  
GSS.  
KEYWORDS  
SOURCE  
ORGANISM  
Canis familiaris (dog)  
Canis familiaris  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
REFERENCE  
1 (bases 1 to 279)  
Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,  
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and  
Venter, J.C.  
The dog genome: survey sequencing and comparative analysis  
Science 301 (5641), 1898-1903 (2003)  
JOURNAL  
MEDLINE  
22875432  
PUBMED  
14512627  
COMMENT  
Contact: Kirkness EF  
The Institute for Genomic Research  
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
Rockville, MD 20850, USA  
Tel: 301-838-0200  
Fax: 301-838-0208  
Email: ekirknes@tigr.org

Class: shotgun.  
 Location/Qualifiers  
 1..279  
 /organism="Canis familiaris"  
 /mol\_type="genomic DNA"  
 /strain="Standard Poodle"  
 /db\_xref="taxon:9615"  
 /clone\_lib="Dog Library"  
 /note="Site 1: BstXI; Libraries were prepared from peripheral blood"

## ORIGIN

Query Match 4.6%; Score 22; DB 29; Length 279;  
 Best Local Similarity 100.0%; Pred. No. 3.8;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 TATTATCTCTGCTTTTGTCT 403  
 |||||  
 Db 182 TATTATCTCTGCTTTTGTCT 203

## RESULT 2

BG359460  
 LOCUS BG359460 373 bp mRNA linear EST 28-NOV-2001  
 DEFINITION sac24e07.y1 Gm-cl051 Glycine max cDNA clone GENOME SYSTEMS CLONE  
 ID: Gm-cl051-3278 5', mRNA sequence.

## ACCESSION

VERSION BG359460

## KEYWORDS

EST.

## SOURCE

Glycine max (soybean)

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

## REFERENCE

1 (bases 1 to 373)

## AUTHORS

Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V.,  
 Khanna,A., Bolla,B., Marra,M., Hillier,B., Kucaba,T., Martin,J.,  
 Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,  
 Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,  
 Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,  
 McCann,R., Waterston,R. and Wilson,R.

## TITLE

Public Soybean EST-Project

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST-Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu

## FEATURES

source

1..373  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl051-3278"  
 /tissue\_type="floral meristematic mRNA"  
 /lab\_host="DH10B"  
 /clone\_lib="Gm-cl051"  
 /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
 XhoI; The cDNA library was constructed from floral  
 meristematic mRNA provided by Dr. Halina Knap of Clemson  
 University. Complementary DNA was synthesized from mRNA  
 using a primer consisting of a poly(dT) sequence with a  
 XhoI restriction site. EcoRI adapters were ligated to the  
 blunt-ended cDNA fragments followed by XhoI digestion. The  
 cDNA fragments were directionally cloned into the  
 EcoRI-XhoI restriction site of the pBluescript vector. The

ligated cDNA fragments were transformed into DH10B host  
 cells (GibcoBRL). This library was constructed in the  
 laboratory of Dr. Randy Shoemaker."

## ORIGIN

Query Match 4.4%; Score 21; DB 12; Length 373;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 367 ATAATTATGTTTCATTATTAT 387  
 |||||  
 Db 213 ATAATTATGTTTCATTATTAT 233

## RESULT 3

AZ756615

## LOCUS

AZ756615

## DEFINITION

ev13h12.r1 PAX3 CASTING Library 'ev' Homo sapiens genomic clone

ev13h12 random, genomic survey sequence.

## ACCESSION

VERSION AZ756615

## KEYWORDS

GSS.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 594)

## AUTHORS

Barber,T.D., Barber,M.C., Tomescu,O., Barr,F.G., Ruben,S. and  
 Friedman,T.B.

## TITLE

Identification of Target Genes Regulated by PAX3 and PAX3--FKHR in

Embryogenesis and Alveolar Rhabdomyosarcoma

Genomics 79 (3), 278-284 (2002)

21853298

PUBMED

11863357

COMMENT

Contact: Friedman TB

Laboratory of Molecular Genetics

National Institutes on Deafness and Other Communication Disorders,

National Institutes of Health

5 Research Court, Room 2A-15, Rockville, MD 20850, USA

Tel: 301 402 7580

Fax: 301 496 7882

Email: friedman@nidcd.nih.gov

Plate: 13 row: h column: 12

Seq primer: -28M13 reverse primer (Amersham)

Class: random plasmid subclone.

Location/Qualifiers

1..594

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/clone="ev13h12"

/sex="Male"

/lab\_host="DH10B"

/note="Vector: pGEM-T Easy; Human genomic DNA was

partially digested with Sau3AI, ligated to ds linkers,

and enriched for binding to human PAX3Q+ protein using a

Whole Genome PCR-based strategy. DNA fragments containing

putative PAX3Q+ binding sites were amplified by PCR and

cloned into pGEM-T Easy (Promega). The ligation products

were transformed into DH10B electrocompetent cells (Life

Technologies)."

## ORIGIN

Query Match 4.4%; Score 21; DB 28; Length 594;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 AGATAAGACTGACCCAGGCC 316  
 |||||  
 Db 189 AGATAAGACTGACCCAGGCC 209

```

RESULT 4
A2756578      .595 bp   DNA       linear   GSS 01-MAR-2001
LOCUS         ev13e08.r1 PAX3 CASTING Library 'ev' Homo sapiens genomic clone
DEFINITION    ev13e08 random, genomic survey sequence.
ACCESSION     A2756578
VERSION       A2756578.1 GI:13176030
KEYWORDS      GSS.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 595)
AUTHORS       Barber,T.D., Barber,M.C., Tomescu,O., Barr,F.G., Ruben,S. and
              Friedman,T.B.
TITLE         Identification of Target Genes Regulated by PAX3 and PAX3--FKHR in
              Embryogenesis and Alveolar Rhabdomyosarcoma
JOURNAL        Genomics 79 (3), 278-284 (2002)
MEDLINE        21853298
PUBMED        11863357
COMMENT       Contact: Friedman TB
              Laboratory of Molecular Genetics
              National Institute on Deafness and Other Communication Disorders,
              National Institutes of Health
              5 Research Court, Room 2A-15, Rockville, MD 20850, USA
              Tel: 301 402 7580
              Fax: 301 496 7882
              Email: friedman@nidcd.nih.gov
              Plate: 13 row: e column: 08
              Seq primer: -28M13 reverse primer (Amersham)
              Class: random plasmid subclone.
              Location/Qualifiers
              1..595
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /clone="ev13e08"
              /sex="Male"
              /lab_host="DH10B"
              /clone_lib="PAX3 CASTING Library 'ev'"
              /note="vector: pGEM-T Easy; Human genomic DNA was
              partially digested with Sau3AI, ligated to ds linkers,
              and enriched for binding to human PAX3DQ+ protein using a
              Whole Genome PCR-based strategy. DNA fragments containing
              putative PAX3DQ+ binding sites were amplified by PCR and
              cloned into pGEM-T Easy (Promega). The ligation products
              were transformed into DH10B electrocompetent cells (Life
              Technologies)."
```

```

FEATURES             source
source               1..595
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /clone="ev13e08"
                    /sex="Male"
                    /lab_host="DH10B"
                    /clone_lib="PAX3 CASTING Library 'ev'"
                    /note="vector: pGEM-T Easy; Human genomic DNA was
                    partially digested with Sau3AI, ligated to ds linkers,
                    and enriched for binding to human PAX3DQ+ protein using a
                    Whole Genome PCR-based strategy. DNA fragments containing
                    putative PAX3DQ+ binding sites were amplified by PCR and
                    cloned into pGEM-T Easy (Promega). The ligation products
                    were transformed into DH10B electrocompetent cells (Life
                    Technologies)."
```

```

ORIGIN
Query Match      4.4%; Score 21; DB 28; Length 595;
Best Local Similarity 100.0%; Pred. NO. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 296 AGATAAGACTGACCCAGGCC 316
    |||||
Db 199 AGATAAGACTGACCCAGGCC 219

RESULT 6
BG483395
LOCUS         BG483395
DEFINITION    602504267F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4617617 5',
              mRNA sequence.
ACCESSION     BG483395
VERSION       BG483395.1 GI:13415674
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 772)
AUTHORS       NIH-MGC http://mgi.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
JOURNAL        Contact: Robert Strausberg, Ph.D.
              Email: cgabbs-r@mail.nih.gov
              Tissue Procurement: CLONTECH Laboratories, Inc.
              cDNA Library Preparation: CLONTECH Laboratories, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/ILNL at:
              http://image.llnl.gov
              Plate: L1CMI374 row: f column: 18
              High quality sequence stop: 658.
              Location/Qualifiers
```

```

FEATURES             source
source               1..772
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /clone="ev13e08"
                    /sex="Male"
                    /lab_host="DH10B"
                    /clone_lib="PAX3 CASTING Library 'ev'"
                    /note="vector: pGEM-T Easy; Human genomic DNA was
                    partially digested with Sau3AI, ligated to ds linkers,
                    and enriched for binding to human PAX3DQ+ protein using a
                    Whole Genome PCR-based strategy. DNA fragments containing
                    putative PAX3DQ+ binding sites were amplified by PCR and
                    cloned into pGEM-T Easy (Promega). The ligation products
                    were transformed into DH10B electrocompetent cells (Life
                    Technologies)."
```

```

ORIGIN
Query Match      4.4%; Score 21; DB 28; Length 595;
Best Local Similarity 100.0%; Pred. NO. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 296 AGATAAGACTGACCCAGGCC 316
    |||||
Db 195 AGATAAGACTGACCCAGGCC 215

RESULT 5
A2756590      602 bp   DNA       linear   GSS 01-MAR-2001
LOCUS         ev13f08.r1 PAX3 CASTING Library 'ev' Homo sapiens genomic clone
DEFINITION    ev13f08 random, genomic survey sequence.
ACCESSION     A2756590
VERSION       A2756590.1 GI:13176042
KEYWORDS      GSS.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 602)
AUTHORS       Barber,T.D., Barber,M.C., Tomescu,O., Barr,F.G., Ruben,S. and
              Friedman,T.B.
```

```

source
1. .772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4617617"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MSC_77"
/Note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgctggcc); Site_2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTATGCCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGAGCGCGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MSC Library."

ORIGIN
Query Match 4.4%; Score 21; DB 12; Length 772;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AAAGACAGCAGCAATGCTG 143
|||||
Db 423 AAAGACAGCAGCAATGCTG 443

RESULT 7
CC969844
LOCUS
DEFINITION
CC969844
ACCESSION
CC969844
VERSION
CC969844.1 GI:33825820
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 826)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: ZUAFU1TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
Location/Qualifiers
1. .826
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBPA0054L01"
/clone_lib="ZM_3.0 4.0 kb"
/Note="Vector: pBGSK-; Site_1: HincII; 3-4 kb 'unfiltered'
genomic DNA library"

ORIGIN
Query Match 4.4%; Score 21; DB 29; Length 826;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 377 TTCATTATTATCTCTGCTTT 397
|||||
Db 15 TTCATTATTATCTCTGCTTT 35

source
1. .772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4617617"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MSC_77"
/Note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgctggcc); Site_2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTATGCCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGAGCGCGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MSC Library."

ORIGIN
Query Match 4.4%; Score 21; DB 12; Length 772;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AAAGACAGCAGCAATGCTG 143
|||||
Db 423 AAAGACAGCAGCAATGCTG 443

RESULT 8
CC949260
LOCUS
DEFINITION
CC949260
ACCESSION
CC949260
VERSION
CC949260.1 GI:10527019
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 196)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
spinal cord tissue cDNA Library Preparation: M.B. Soares Lab Clone
distribution: Researchers may obtain BMAP cDNA clones from RESEARCH
GENETICS. It should be noted that Bento Soares is generating a
small number of additional specialized non-redundant arrays of BMAP
cDNAs whose availability will be considered under appropriate and
limited collaborative arrangements
Seq primer: M13 Forward
POLYA=yes.

FEATURES
Location/Qualifiers
1. .196
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH3-ave-h-11-0-UI"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH_BMAP_M_S4"
/Note="Vector: pTV3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NIH_BMAP_M_S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH_BMAP_M_S4,
NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
(NIH_BMAP_M_S4) was constructed as follows: PCR amplified
cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)

```



was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (lifestyle technologies) to generate the NIH BMAP M.S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)  
 TAG\_TISSUE=spinal-cord  
 TAG\_LIB=NIH\_BMAP\_M\_S4  
 TAG\_SEQ=TCAA"

## ORIGIN

Query Match 4.2%; Score 20; DB 10; Length 196;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 369 AATTATGTTTCATTATTATT 388  
 Db 40 AATTATGTTTCATTATTATT 21

## RESULT 9

AA759629/c  
 LOCUS 216 bp mRNA linear EST 23-JAN-1998  
 DEFINITION W56506.r1 Soares\_mammary\_gland\_NMLMG Mus musculus cDNA clone  
 IMAGE:1247795 5', mRNA sequence.  
 ACCESSION AA759629  
 VERSION AA759629.1 GI:2807423  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

## REFERENCE

1 (bases 1 to 216)  
 Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisli, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

## TITLE

The WashU-HHMI Mouse EST Project

## JOURNAL

Unpublished (1996)

## COMMENT

Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:661483

Seq primer: -28m13 rev2 ET from Amersham

High quality sequence stop: 192.

Location/Qualifiers

1..216

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="IMAGE:1247795"

/sex="female (lactating)"

/tissue type="mammary gland"

/lab host="DH10B"

/clone lib="Soares mammary gland NMLMG"

/note="Vector: p7T3D-Pac (Pharmacia) with a modified

polylinker; 1st strand cDNA was prepared from mammary

gland tissue from a lactating female, and was then primed

with a Not I - oligo(dT) primer. Double-stranded cDNA was

ligated to Eco RI adaptors (Pharmacia), digested with Not

I and cloned into the Not I and Eco RI sites of the

modified p7T73 vector. Library is normalized. Library

was constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 4.2%; Score 20; DB 9; Length 216;

Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 369 AATTATGTTTCATTATTATT 388  
 Db 213 AATTATGTTTCATTATTATT 194

## RESULT 10

AW913044/c  
 LOCUS 247 bp mRNA linear EST 25-MAY-2000  
 DEFINITION uf48d01.y1 Soares\_mammary\_gland\_NMLMG Mus musculus cDNA clone  
 IMAGE:1514593 5', mRNA sequence.  
 ACCESSION AW913044  
 VERSION AW913044.1 GI:8078681  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

## REFERENCE

1 (bases 1 to 247)  
 Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Eukaryota; Murinae; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)

## AUTHORS

Other ESTs: uf48d01.x1

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:941445

Seq primer: -40RP from Gibco

High quality sequence stop: 241.

Location/Qualifiers

1..247

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="IMAGE:1514593"

/sex="female (lactating)"

/tissue type="mammary gland"

/lab host="DH10B"

/clone lib="Soares mammary gland NMLMG"

/note="Vector: p7T3D-Pac (Pharmacia) with a modified

polylinker; 1st strand cDNA was prepared from mammary

gland tissue from a lactating female, and was then primed

with a Not I - oligo(dT) primer. Double-stranded cDNA was

ligated to Eco RI adaptors (Pharmacia), digested with Not

I and cloned into the Not I and Eco RI sites of the

modified p7T73 vector. Library is normalized. Library

was constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 4.2%; Score 20; DB 10; Length 247;  
 Best Local Similarity 100.0%; Pred. No. 41;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 AATTATGTTTCATTATTATT 388

Db 208 AATTATGTTTCATTATTATT 189

## RESULT 11

T38551  
 LOCUS 287 bp mRNA linear EST 11-JAN-1995  
 DEFINITION EST104054 S. cerevisiae strain X2180-1A Saccharomyces cerevisiae  
 cDNA 3' end, mRNA sequence.  
 ACCESSION T38551  
 VERSION T38551.1 GI:622368  
 KEYWORDS EST.  
 SOURCE Saccharomyces cerevisiae (baker's yeast)  
 ORGANISM Saccharomyces cerevisiae

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

1 (bases 1 to 287)

Weinstock, K.

Saccharomycetes cerevisiae cDNAs

Unpublished (1995)

Contact: Weinstock, K. and Venter, J.C.

The Institute for Genomic Research

932 Clopper Rd, Gaithersburg, MD 20878

Tel: 3018699056

Fax: 3018699423

Email: tdbinfo@tdb.tigr.org

For clone availability please contact the TIGR Database

(tdbinfo@tdb.tigr.org)

Seq primer: M13-21.

Location/Qualifiers

1..287

/organism="Saccharomycetes cerevisiae"

/mol\_type="mRNA"

/strain="X2180-1A"

/db\_xref="taxon:4932"

/clone\_lib="S. cerevisiae strain X2180-1A"

/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:

XhoI"

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/clone="UI-R-CAL-bja-n-15-0-UI"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="UI-R-CAL"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CAL

library is a subtracted library derived from the following

tissues: thalamus, cerebellum, hypothalamus, medulla, testis,

pons, midbrain, cerebral cortex, corpus striatum, and

hippocampus. For a detailed description of the

library from which this clone was derived, please visit

our web site at [ratest.eng.uiowa.edu](http://ratest.eng.uiowa.edu). The subtraction

has been previously described in (Bonaldo, Lennon and

Soares, Genome Research 6:791-806, 1996)

TAG\_RISUE=hippocampus

TAG\_LIB=UI-R-CAL

TAG\_SEQ=GATTG"



```

FEATURES             Seq primer: M13 Forward.
source              Location/Qualifiers
1..331
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="C57BL/6J"
    /db_xref="AtCC (inhost):1363847"
    /db_xref="taxon:10090"
    /clone="C0013C12"
    /sex="unknown"
    /tissue_type="ectoplacental cone"
    /dev_stage="embryonic day 7.5 postconception"
    /lab_host="DH10B"
    /clone_lib="Mouse 7.5 dpc embryo ectoplacental cone cDNA
    library"
    /note="Organ: embryo; Vector: pSPORT1 (Life Technologies);
    Site_1: SalI; Site_2: NotI; Total RNAs were extracted from
    ectoplacental cone of 7.5-dpc embryos. The
    double-stranded cDNA was synthesized from total RNAs with
    an Oligo(dt) primer. The library was constructed by Minoru
    S. H. Ko."

```

ORIGIN

```

Query Match          4.2%; Score 20; DB 9; Length 331;
Best Local Similarity 100.0%; Pred.No. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 AATTATGTTTCATTATTATT 388
    |||||
Db 281 AATTATGTTTCATTATTATT 300

```

Search completed: July 31, 2004, 10:50:08  
 Job time : 2005 secs

JanCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 09:32:50 ; Search time 336 Seconds  
(without alignments)

6946.115 Million cell updates/sec

Title: US-10-020-540A-1

Perfect score: 476

Sequence: 1 aactatgctgacgaagat.....aactccctgcttacaacc 476

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 3222919 seqs, 2451570024 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:
- 2: /cgn2\_6/ptodata/2/pubpna/PCT NEW PUB.seq:
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW PUB.seq:
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW PUB.seq:
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW PUB.seq:
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW PUB.seq:
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:
- 14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:
- 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:
- 16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:
- 17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW PUB.seq:
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW PUB.seq:
- 19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	476	100.0	476	US-10-020-540A-1	Sequence 1, Appli
2	476	100.0	518	US-10-020-540A-2	Sequence 2, Appli
3	369	77.5	651	US-10-020-540A-3	Sequence 3, Appli
4	369	77.5	701	US-10-020-540A-4	Sequence 4, Appli
5	21	4.4	373	US-10-024-599-104814	Sequence 104814,
6	21	4.4	622	US-10-027-632-229768	Sequence 229768,
7	21	4.4	622	US-10-027-632-229768	Sequence 229768,
8	20	4.2	99973	US-10-085-117-109	Sequence 109, App
9	19	4.0	1553	US-10-437-963-22143	Sequence 22143, A
10	19	4.0	3237	US-10-437-963-22139	Sequence 22139, A
11	19	4.0	5675	US-10-369-493-29305	Sequence 29305, A
12	18	3.8	28	US-10-020-540A-10	Sequence 10, Appl
13	18	3.8	327	US-09-867-701-603	Sequence 603, App
14	18	3.8	402	US-10-437-963-76846	Sequence 76846, A

15	18	3.8	454	9	US-09-867-701-949	Sequence 949, App
16	18	3.8	536	9	US-09-867-701-733	Sequence 733, App
c 17	18	3.8	553	13	US-10-424-599-41806	Sequence 41806, A
c 18	18	3.8	570	13	US-10-027-632-184726	Sequence 184726,
c 19	18	3.8	570	16	US-10-027-632-184726	Sequence 184726,
20	18	3.8	637	13	US-10-027-632-211784	Sequence 211784,
21	18	3.8	637	16	US-10-027-632-211784	Sequence 211784,
22	18	3.8	668	13	US-10-027-632-260421	Sequence 260421,
23	18	3.8	668	16	US-10-027-632-260421	Sequence 260421,
c 24	18	3.8	872	15	US-10-029-386-22974	Sequence 22974, A
c 25	18	3.8	1752	17	US-10-437-963-54963	Sequence 54963, A
c 26	18	3.8	2259	13	US-10-027-632-255809	Sequence 255809,
c 27	18	3.8	2259	16	US-10-027-632-255810	Sequence 255810,
c 28	18	3.8	2259	16	US-10-027-632-255809	Sequence 255809,
c 29	18	3.8	2259	16	US-10-027-632-255810	Sequence 255810,
c 30	18	3.8	2597	15	US-10-037-270-48	Sequence 48, Appl
c 31	18	3.8	2597	16	US-10-117-722-48	Sequence 48, Appl
c 32	18	3.8	7380	13	US-09-939-853A-10	Sequence 10, Appl
c 33	18	3.8	7380	13	US-09-939-853A-12	Sequence 12, Appl
c 34	18	3.8	16001	15	US-10-085-959-21	Sequence 21, Appl
c 35	18	3.8	38584	15	US-10-114-170-50	Sequence 50, Appl
c 36	18	3.8	63686	13	US-10-087-192-466	Sequence 466, App
c 37	18	3.8	79799	13	US-10-087-192-346	Sequence 346, App
c 38	18	3.8	138115	17	US-10-322-281-377	Sequence 377, App
c 39	17	3.6	162	13	US-10-085-783A-49002	Sequence 49002, A
c 40	17	3.6	182	16	US-10-242-535A-49002	Sequence 49002, A
c 41	17	3.6	231	10	US-09-535-459-2040	Sequence 2040, Ap
c 42	17	3.6	247	9	US-09-983-965-2222	Sequence 2222, Ap
c 43	17	3.6	270	10	US-09-535-459-2044	Sequence 2044, Ap
c 44	17	3.6	295	9	US-09-294-093B-233	Sequence 233, App
c 45	17	3.6	302	10	US-09-535-459-1934	Sequence 1934, Ap

ALIGNMENTS

RESULT 1

US-10-020-540A-1  
; Sequence 1, Application US/10020540A  
; Publication No. US20020182593A1  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Gusui  
; APPLICANT: McBrady, Kevin  
; TITLE OF INVENTION: Strawberry Vein Banding Virus Promoter  
; TITLE OF INVENTION: (SVBV) Promoter  
; FILE REFERENCE: 0173.210us  
; CURRENT APPLICATION NUMBER: US/10/020,540A  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/245,354  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 476  
; TYPE: DNA  
; ORGANISM: Strawberry Vein Banding Virus (SVBV) Strain E3  
US-10-020-540A-1

Query Match	100.0%	Score 476;	DB 14;	Length 476;
Best Local Similarity	100.0%;	Pred. No. 7.1e-240;		
Matches 476;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AACATATGCTGATGACAGATAATCTTAATAGCAATTAATTCAGAAATTAATCAAGGAGAAA	60	
Db	1	AACATATGCTGATGACAGATAATCTTAATAGCAATTAATTCAGAAATTAATCAAGGAGAAA	60	
Qy	61	GAATTAATACTCTTTTTCAGAAATTAATGAGCCGCTTTTCAAGTGGCCAGCTAGTATCACT	120	
Db	61	GAATTAATACTCTTTTTCAGAAATTAATGAGCCGCTTTTCAAGTGGCCAGCTAGTATCACT	120	
Qy	121	GAAGAAGCAGCAGACAAATGCTGCTCGATGACACAGAACCAATCTTTGCACGATGT	180	
Db	121	GAAGAAGCAGCAGACAAATGCTGCTCGATGACACAGAACCAATCTTTGCACGATGT	180	

QY 181 GAAGCGAGCAGAGTGTCCACAGAGCGAGCTCAGAAAAGGCGATCTTCTACCGACACAGAA 240  
Db 181 GAAGCGAGCAGAGTGTCCACAGAGCGAGCTCAGAAAAGGCGATCTTCTACCGACACAGAA 240  
QY 241 AAAGACAACACACAGCTCATCATCCAAACATGTAGACTGTGCTTATGCGTCCGCTCAAGATA 300  
Db 241 AAAGACAACACACAGCTCATCATCCAAACATGTAGACTGTGCTTATGCGTCCGCTCAAGATA 300  
QY 301 AGACTGACCCAGGCGCAGCAGCTAAAGAGAAATATGCAAGTGGTCTTAGCTCCACTTTA 360  
Db 301 AGACTGACCCAGGCGCAGCAGCTAAAGAGAAATATGCAAGTGGTCTTAGCTCCACTTTA 360  
QY 361 GCTTTAATAATATGTTTTCATTATTTATTTCTGCTGCTTTTCTCTCTATATAAAGAGCTTGT 420  
Db 361 GCTTTAATAATATGTTTTCATTATTTATTTCTGCTGCTTTTCTCTCTATATAAAGAGCTTGT 420  
QY 421 ATTTCATTGAGCGAGGCGGACACACACACAGAGACCTCCCTGCTTACAAACC 476  
Db 421 ATTTCATTGAGCGAGGCGGACACACACACAGAGACCTCCCTGCTTACAAACC 476

## RESULT 2

US-10-020-540A-2

; Sequence 2, Application US/10020540A

; Publication No. US20020182593A1

; GENERAL INFORMATION:

; APPLICANT: Wu, Gusui

; APPLICANT: McBryde, Kevin

; TITLE OF INVENTION: Strawberry Vein Banding Virus Promoter

; FILE REFERENCE: 0173.210us

; CURRENT APPLICATION NUMBER: US/10/020,540A

; CURRENT FILING DATE: 2001-10-30

; PRIOR APPLICATION NUMBER: 60/245,354

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 518

; TYPE: DNA

; ORGANISM: Strawberry Vein Banding Virus (SVBV) Strain E3

US-10-020-540A-2

Query Match 100.0%; Score 476; DB 14; Length 518;  
Best Local Similarity 100.0%; Pred. No. 7.2e-240;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTATGCTGATCAGACAGATATTTCTAATAGCAATTTATTCAGAAATTAATCAGGAGAA 60  
Db 15 AACTATGCTGATCAGACAGATATTTCTAATAGCAATTTATTCAGAAATTAATCAGGAGAA 74  
QY 61 GAATTAATAACTCTTTTCAAGATATGAAGCCGCTTTTCAAGTGCCAGCTAGCTATCACT 120  
Db 75 GAATTAATAACTCTTTTCAAGATATGAAGCCGCTTTTCAAGTGCCAGCTAGCTATCACT 134  
QY 121 GAAAGACGACGACAGATGCTTCCATGACACAGACACACATCTTTCGACGAGATGT 180  
Db 135 GAAAGACGACGACAGATGCTTCCATGACACAGACACACATCTTTCGACGAGATGT 194  
QY 181 GAAGCAGCAGAGTGTGTCACAGAGCGCAGCTCAGAAAAGGCGATCTTCTACCGACACAGAA 240  
Db 195 GAAGCAGCAGAGTGTGTCACAGAGCGCAGCTCAGAAAAGGCGATCTTCTACCGACACAGAA 254  
QY 241 AAAGACAACACACAGCTCATCATCCAAACATGTAGACTGTGCTTATGCGTCCGCTCAAGATA 300  
Db 255 AAAGACAACACACAGCTCATCATCCAAACATGTAGACTGTGCTTATGCGTCCGCTCAAGATA 314  
QY 301 AGACTGACCCAGGCGCAGCAGCTAAAGAGAAATATGCAAGTGGTCTTAGCTCCACTTTA 360  
Db 315 AGACTGACCCAGGCGCAGCAGCTAAAGAGAAATATGCAAGTGGTCTTAGCTCCACTTTA 374  
QY 361 GCTTTAATAATATGTTTTCATTATTTATTTCTGCTGCTTTTCTCTATATAAAGAGCTTGT 420

Db 375 GCTTTAATAATATGTTTTCATTATTTATTTCTGCTGCTTTTCTCTATATAAAGAGCTTGT 434  
QY 421 ATTTCATTGAGCGAGGCGGACACACACAGAGACCTCCCTGCTTACAAACC 476  
Db 435 ATTTCATTGAGCGAGGCGGACACACACAGAGACCTCCCTGCTTACAAACC 490

## RESULT 3

US-10-020-540A-3

; Sequence 3, Application US/10020540A

; Publication No. US20020182593A1

; GENERAL INFORMATION:

; APPLICANT: Wu, Gusui

; APPLICANT: McBryde, Kevin

; TITLE OF INVENTION: Strawberry Vein Banding Virus Promoter

; FILE REFERENCE: 0173.210us

; CURRENT APPLICATION NUMBER: US/10/020,540A

; CURRENT FILING DATE: 2001-10-30

; PRIOR APPLICATION NUMBER: 60/245,354

; PRIOR FILING DATE: 2000-11-01

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 651

; TYPE: DNA

; ORGANISM: Strawberry Vein Banding Virus (SVBV) Strain E3

US-10-020-540A-3

Query Match 77.5%; Score 369; DB 14; Length 651;  
Best Local Similarity 100.0%; Pred. No. 1.6e-183;  
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 GCTAGCTATCATCTGAAAGACAGCAAGCAATGTTGTTCTGATGCGACCAAGACCATCT 167  
Db 278 GCTAGCTATCATCTGAAAGACAGCAAGCAATGTTGTTCTGATGCGACCAAGACCATCT 337  
QY 168 TTGCAGCAGATGTGAAGCAGCGCAGAGTGGTCCACAGACGACCTCAGAAAAGGCATCTTC 227  
Db 338 TTGCAGCAGATGTGAAGCAGCGCAGAGTGGTCCACAGACGACCTCAGAAAAGGCATCTTC 397  
QY 228 TACCGACACAGAAAAGACCAACCAAGCTCATCATTCACCATGTAGACTGTCTGTATGCG 287  
Db 398 TACCGACACAGAAAAGACCAACCAAGCTCATCATTCACCATGTAGACTGTCTGTATGCG 457  
QY 268 TCGGCTGAAGATAAGACTGACCCCGCAGCAGCTAAAGAGAAATAATGCAAGTGGTCC 347  
Db 458 TCGGCTGAAGATAAGACTGACCCCGCAGCAGCTAAAGAGAAATAATGCAAGTGGTCC 517  
QY 348 TAGCTCCACTTTAGCTTTAATAATATGTTTCAATTTATTTCTGCTTTTGTCTCTAT 407  
Db 518 TAGCTCCACTTTAGCTTTAATAATATGTTTCAATTTATTTCTGCTTTTGTCTCTAT 577  
QY 408 ATAAAGAGCTTGTATTTTTCATTGAGGCGAGGCGGACACACACAGAACTCCCTGCG 467  
Db 578 ATAAAGAGCTTGTATTTTTCATTGAGGCGAGGCGGACACACACAGAACTCCCTGCG 637  
QY 468 TTACAAACC 476  
Db 638 TTACAAACC 646

## RESULT 4

US-10-020-540A-4

; Sequence 4, Application US/10020540A

; Publication No. US20020182593A1

; GENERAL INFORMATION:

; APPLICANT: Wu, Gusui

; APPLICANT: McBryde, Kevin

; TITLE OF INVENTION: Strawberry Vein Banding Virus Promoter

; FILE REFERENCE: 0173.210us

CURRENT APPLICATION NUMBER: US/10/020,540A  
CURRENT FILING DATE: 2001-10-30  
PRIOR APPLICATION NUMBER: 60/245,354  
PRIOR FILING DATE: 2000-11-01  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 701  
TYPE: DNA  
ORGANISM: Strawberry Vein Banding Virus (SVBV) Strain E3  
US-10-020-540A-4

Query Match 77.5%; Score 369; DB 14; Length 701;  
Best Local Similarity 100.0%; Pred. No. 1.6e-183; Indels 0; Gaps 0;  
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 GCTAGCTATCACTGAAAGACAGACAGACAATGGTCTCGATGACCAAGACCACTCT 167  
DB 319 GCTAGCTATCACTGAAAGACAGACAGACAATGGTCTCGATGACCAAGACCACTCT 378  
QY 168 TTGCAGCAGATGTGAAGCGCGAGGTGTCACAGAGCGCACTCAGAAAAGGCATCTTC 227  
DB 379 TTGCAGCAGATGTGAAGCGCGAGGTGTCACAGAGCGCACTCAGAAAAGGCATCTTC 438  
QY 228 TACCGACAGAAAAAGACAAACACAGCTCATCTCAACATGTAGACTGTGTTATGGG 287  
DB 439 TACCGACAGAAAAAGACAAACACAGCTCATCTCAACATGTAGACTGTGTTATGGG 498  
QY 288 TCGGCTGAAGATAAGACTGACCCAGCCAGCACTAAAGAGAAATATGCAAGTGTCC 347  
DB 499 TCGGCTGAAGATAAGACTGACCCAGCCAGCACTAAAGAGAAATATGCAAGTGTCC 558  
QY 348 TAGCTCCACTTTAGCTTTAATATTATTTTCAATATTATTCCTCTCTCTCTCTAT 407  
DB 559 TAGCTCCACTTTAGCTTTAATATTATTTTCAATATTATTCCTCTCTCTCTCTAT 618  
QY 408 ATAAAGAGCTTGTATTTTTCATTTGAAGGCGAGGCGGAACACACACAGAACTCCCTGC 467  
DB 619 ATAAAGAGCTTGTATTTTTCATTTGAAGGCGAGGCGGAACACACACAGAACTCCCTGC 678  
QY 468 TTACAAACC 476  
DB 679 TTACAAACC 687

RESULT 5  
US-10-424-599-104814  
Sequence 104814, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 104814  
LENGTH: 373  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_65663C.1  
US-10-424-599-104814

Query Match 4.4%; Score 21; DB 13; Length 373;  
Best Local Similarity 100.0%; Pred. No. 2.4; Indels 0; Gaps 0;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 367 ATAATTATGTTTCATTATTAT 387

DB 213 ATAATTATGTTTCATTATTAT 233

RESULT 6  
US-10-027-632-229768/c  
Sequence 229768, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 229768  
LENGTH: 622  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-229768

Query Match 4.4%; Score 21; DB 13; Length 622;  
Best Local Similarity 100.0%; Pred. No. 2.5; Indels 0; Gaps 0;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 TAAAGAGAAATATGCAAGT 342  
DB 545 TAAAGAGAAATATGCAAGT 525

RESULT 7  
US-10-027-632-229768/c  
Sequence 229768, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 229768
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-229768

Query Match
Best Local Similarity 4.4%; Score 21; DB 16; Length 622;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 TAAGAGAGAAATAATGCAAGT 342
Db 545 TAAGAGAGAAATAATGCAAGT 525

RESULT 8
US-10-085-117-109/c
; Sequence 109, Application US/10085117
; Publication No. US20030232334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 109
; LENGTH: 99973
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: Variation
; LOCATION: (1)...(99973)
; OTHER INFORMATION: n = any nucleotide
US-10-085-117-109

Query Match
Best Local Similarity 4.2%; Score 20; DB 16; Length 99973;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 ATTTTCATTGAAGGCAGAG 440
Db 76871 ATTTTCATTGAAGGCAGAG 76852

RESULT 9
US-10-437-963-22143/c
; Sequence 22143, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: RICE NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 22143
; LENGTH: 1553
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_27347C.1
US-10-437-963-22143

Query Match
Best Local Similarity 4.0%; Score 19; DB 17; Length 3237;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 CAGAACACATCTTTGCAG 173
Db 2536 CAGAACACATCTTTGCAG 2518

RESULT 11
US-10-369-493-29305/c
; Sequence 29305, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 29305
; LENGTH: 5675
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_27343C.1
US-10-369-493-29305

Query Match
Best Local Similarity 4.0%; Score 19; DB 16; Length 5675;
```



Best Local Similarity 100.0%; Pred. No. 35;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 AGAGGCGACACACACACA 455  
Db 3868 AGAGGCGACACACACACA 3850

RESULT 12

US-10-020-540A-10/c  
; Sequence 10, Application US/10020540A  
; Publication No. US20020182593A1  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Gusui  
; APPLICANT: McBryce, Kevin  
; TITLE OF INVENTION: Strawberry Vein Banding Virus Promoter  
; TITLE OF INVENTION: (SVBV) Promoter  
; FILE REFERENCE: 0173.210us  
; CURRENT APPLICATION NUMBER: US/10/020,540A  
; CURRENT FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/245,354  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 28  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Primer  
US-10-020-540A-10

Query Match 3.8%; Score 18; DB 14; Length 28;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 459 CCTCCCTGCTTACAAACC 476  
Db 28 CCTCCCTGCTTACAAACC 11

RESULT 13

US-09-867-701-603/c  
; Sequence 603, Application US/09867701  
; Patent No. US20020132237A1  
; GENERAL INFORMATION:  
; APPLICANT: Aglate, Paul A.  
; APPLICANT: Jones, Robert  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER  
; FILE REFERENCE: 210121.497  
; CURRENT APPLICATION NUMBER: US/09/867,701  
; CURRENT FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 603  
; LENGTH: 327  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(327)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-867-701-603

Query Match 3.8%; Score 18; DB 9; Length 327;  
Best Local Similarity 100.0%; Pred. No. 89;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 CAGAATATGAAGCCCGCT 94  
Db 153 CAGAATATGAAGCCCGCT 136

RESULT 14

US-10-437-963-76846/c  
; Sequence 76846, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 76846  
; LENGTH: 402  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_76802C.1  
US-10-437-963-76846

Query Match 3.8%; Score 18; DB 17; Length 402;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 GTGAAGCAGCCAGAGTGG 196  
Db 259 GTGAAGCAGCCAGAGTGG 242

RESULT 15

US-09-867-701-949  
; Sequence 949, Application US/09867701  
; Patent No. US20020132237A1  
; GENERAL INFORMATION:  
; APPLICANT: Aglate, Paul A.  
; APPLICANT: Jones, Robert  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER  
; FILE REFERENCE: 210121.497  
; CURRENT APPLICATION NUMBER: US/09/867,701  
; CURRENT FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 949  
; LENGTH: 454  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(454)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-867-701-949

Query Match 3.8%; Score 18; DB 9; Length 454;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 CAGAATATGAAGCCCGCT 94  
Db 140 CAGAATATGAAGCCCGCT 157

Search completed: July 31, 2004, 10:57:04

Mon Aug 2 10:06:59 2004

us-10-020-540a-1.olig.rnpb

Page 6

Job time : 338 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 09:31:35 ; Search time 65 Seconds  
(without alignments)  
4063.948 Million cell updates/sec

Title: US-10-020-540A-1  
Perfect score: 476  
Sequence: 1 aactatgctgatgacaagat.....aacctccctgcttacaacc 476

Scoring table: OLIGO\_NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq:  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	18	3.8	2597	4	US-09-620-312D-48
C 2	18	3.8	38584	4	US-09-453-702B-50
C 3	17	3.6	501	4	US-09-543-681A-597
C 4	17	3.6	528	4	US-09-401-064-275
C 5	17	3.6	726	4	US-09-107-532A-1355
C 6	17	3.6	1566	4	US-09-489-039A-947
C 7	17	3.6	1975	4	US-09-308-345A-6
C 8	17	3.6	3562	3	US-09-360-197-1
C 9	17	3.6	4105	4	US-09-589-567-1
C 10	17	3.6	4617	4	US-09-930-055A-1
C 11	17	3.6	5109	4	US-09-930-055A-2
C 12	16	3.4	298	4	US-09-313-294A-4713
C 13	16	3.4	440	4	US-09-621-976-2230
C 14	16	3.4	813	4	US-09-134-001C-2724
C 15	16	3.4	825	4	US-09-956-171E-1111
C 16	16	3.4	966	4	US-09-134-001C-53
C 17	16	3.4	1113	3	US-09-172-353-1
C 18	16	3.4	1113	4	US-09-793-955-1
C 19	16	3.4	1116	3	US-09-776-971-139
C 20	16	3.4	1221	4	US-09-134-001C-365
C 21	16	3.4	1591	3	US-08-468-856B-9
C 22	16	3.4	1591	3	US-08-468-856B-9
C 23	16	3.4	2013	4	US-09-134-000C-1201
C 24	16	3.4	2813	4	US-09-689-255C-3
C 25	16	3.4	3043	3	US-09-008-271A-14
C 26	16	3.4	3332	4	US-09-448-806C-1
C 27	16	3.4	3401	4	US-09-553-554-3

C 28	16	3.4	3441	4	US-09-095-881-1	Sequence 1, Appli
C 29	16	3.4	3502	2	US-08-724-394A-16	Sequence 16, Appli
C 30	16	3.4	4080	4	US-09-016-434-1292	Sequence 1292, App
C 31	16	3.4	4848	4	US-08-961-527-185	Sequence 185, App
C 32	16	3.4	10952	1	US-08-602-036A-1	Sequence 1, Appli
C 33	16	3.4	10952	2	US-08-502-374A-1	Sequence 1, Appli
C 34	16	3.4	10952	2	US-08-642-407A-1	Sequence 1, Appli
C 35	16	3.4	24979	2	US-08-147-777-3	Sequence 3, Appli
C 36	16	3.4	24979	3	US-08-452-872-3	Sequence 3, Appli
C 37	16	3.4	24979	5	PCT-US93-03985-3	Sequence 3, Appli
C 38	16	3.4	152331	3	US-09-128-155-16	Sequence 16, Appli
C 39	16	3.4	176373	3	US-09-128-155-17	Sequence 17, Appli
C 40	16	3.4	640681	4	US-09-730-988-1	Sequence 1, Appli
C 41	16	3.4	640681	4	US-09-750-988-1	Sequence 1, Appli
C 42	16	3.4	1230025	4	US-09-198-452A-1	Sequence 1, Appli
C 43	16	3.4	1664976	4	US-08-916-421B-1	Sequence 1, Appli
C 44	15	3.2	88	6	5198347-12	Patent No. 5198347
C 45	15	3.2	189	4	US-09-134-000C-2505	Sequence 2505, Ap

ALIGNMENTS

RESULT 1  
US-09-620-312D-48/c  
; Sequence 48, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyen  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunqing  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_Genes Version 1.0  
; SEQ ID NO 48  
; LENGTH: 2597  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (592)..(1272)  
US-09-620-312D-48

Query Match 3.8%; Score 18; DB 4; Length 2597;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 77 CAGAAATATGAAGCCGCT 94  
Db 1658 CAGAAATATGAAGCCGCT 1641

```
RESULT 2
US-09-453-702B-50
; Sequence 50, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
;          Burlaud, Valerie
;          Perna, Nicole T.
;          Plunkett, Guy
;          Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-Dec-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38584
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-09-453-702B-50
Query Match 3.8%; Score 18; DB 4; Length 38584;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 CAGAATATGAGCCCGCT 94
Db 3180 CAGAATATGAGCCCGCT 3197
|||||
RESULT 3
US-09-543-681A-597/c
; Sequence 597, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 597
```

```
; LENGTH: 501
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-597
Query Match 3.6%; Score 17; DB 4; Length 501;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 408 ATAAGAGCTTGTTATT 424
Db 106 ATAAGAGCTTGTTATT 90
|||||
RESULT 4
US-09-401-064-275/c
; Sequence 275, Application US/09401064
; Patent No. 6623923
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeline Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C2
; CURRENT APPLICATION NUMBER: US/09/401,064
; CURRENT FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 371
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 275
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(528)
; OTHER INFORMATION: n = A,T,C or G
US-09-401-064-275
Query Match 3.6%; Score 17; DB 4; Length 528;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AAAGACAGCAAGCAAT 139
Db 391 AAAGACAGCAAGCAAT 375
|||||
RESULT 5
US-09-107-532A-1355
; Sequence 1355, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneké
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1355:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 726 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (B) LOCATION 1...726
; SEQUENCE DESCRIPTION: SEQ ID NO: 1355:
US-09-107-532A-1355

```

```

; Query Match 3.6%; Score 17; DB 4; Length 726;
; Best Local Similarity 100.0%; Pred. No. 38;
; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 325 AGAAGAAATAATCAAG 341
Db 18 AGAAGAAATAATCAAG 34

```

```

RESULT 6
US-09-489-039A-947
; Sequence 947, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR FILING DATE: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 947
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-947

```

```

; Query Match 3.6%; Score 17; DB 4; Length 1566;
; Best Local Similarity 100.0%; Pred. No. 39;
; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 337 GCAAGTGGCTCTAGTCT 353
Db 348 GCAAGTGGCTCTAGTCT 364

```

```

RESULT 7
US-09-308-345A-6
; Sequence 6, Application US/09308345A

```

```

; Patent No. 6569665
; GENERAL INFORMATION:
; APPLICANT: BOEHM, Thomas;
; APPLICANT: DEAR, Neil T. 6569665el calpains, their preparation and use
; TITLE OF INVENTION: NO. 6569665el calpains, their preparation and use
; FILE REFERENCE: 0050/47576
; CURRENT APPLICATION NUMBER: US/09/308,345A
; CURRENT FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: WordPerfect v. 6.1
; SEQ ID NO 6
; LENGTH: 1975
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 44..1963
US-09-308-345A-6

```

```

; Query Match 3.6%; Score 17; DB 4; Length 1975;
; Best Local Similarity 100.0%; Pred. No. 40;
; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 216 AAAGGCATCTTCTACCG 232
Db 1736 AAAGGCATCTTCTACCG 1752

```

```

RESULT 8
US-09-360-197-1
; Sequence 1, Application US/09360197
; Patent No. 6287859
; GENERAL INFORMATION:
; APPLICANT: Basilana, Frederic
; APPLICANT: Lazdunski, Michel
; APPLICANT: Waldmann, Rainer
; APPLICANT: Dweille, Jan R.
; TITLE OF INVENTION: Human and Rat Families of Neuronal Acid-Sensitive
; TITLE OF INVENTION: Cationic Channels, Their Cloning and Applications
; FILE REFERENCE: 989.6706P
; CURRENT APPLICATION NUMBER: US/09/360,197-
; CURRENT FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: 09/129,758
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/095,408
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3562
; TYPE: DNA
; ORGANISM: rattus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (123)..(1700)
US-09-360-197-1

```

```

; Query Match 3.6%; Score 17; DB 3; Length 3562;
; Best Local Similarity 100.0%; Pred. No. 40;
; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 445 ACACACACACAGAACT 461
Db 36 ACACACACACAGAACT 52

```

```

RESULT 9
US-09-589-567-1/c
; Sequence 1, Application US/09589567
; Patent No. 6479730
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Shi, Jinrui

```

;; TITLE OF INVENTION: Maize DNA Ligase II Orthologue and Uses  
;; FILE OF INVENTION: Thereof  
;; FILE REFERENCE: 1125  
;; CURRENT APPLICATION NUMBER: US/09/589,567  
;; PRIOR FILING DATE: 2000-06-07  
;; PRIOR APPLICATION NUMBER: US 60/145,911  
;; PRIOR FILING DATE: 1999-07-27  
;; NUMBER OF SEQ ID NOS: 3  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 1  
;; LENGTH: 4105  
;; TYPE: DNA  
;; ORGANISM: Zea Mays  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (148)...(3801)  
;; NAME/KEY: misc feature  
;; LOCATION: (2713)...(2745)  
;; OTHER INFORMATION: ATP-dependent DNA ligase signature sequence  
US-09-589-567-1

Query Match 3.6%; Score 17; DB 4; Length 4105;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 AACCATCTTTGCAGC 174  
Db 915 AACCATCTTTGCAGC 899

RESULT 10  
US-08-930-055A-1  
; Sequence 1, Application US/08930055A  
; Patent No. 6403561  
; GENERAL INFORMATION:  
; APPLICANT: ROSE, CHRISTINANE; VARGAS,  
; APPLICANT: FROYLAN; BOURGAT, PIERRE; SCHWARTZ, JEAN-  
; APPLICANT: CHARLES; BISHOP, PAUL; BAMBAL, RAMESH;  
; APPLICANT: GANELLIN, CHARON; LEBLOND, BERTRAND; MOORE,  
; APPLICANT: ANDREW; CHAN, SUZANNE; ZHAO, LIHUA.  
; TITLE OF INVENTION: TRIPEPTIDYLPEPTIDASE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIERMAN, MUSERLIAN AND LUCAS  
; ADDRESS: LLP  
; STREET: 600 THIRD AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MSDOS  
; SOFTWARE: MICROSOFT WORD 97  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/930,055A  
; FILING DATE: 19-FEB-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR96/00700  
; FILING DATE: 9-MAY-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR95/05489  
; FILING DATE: 9-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CHARLES A. MUSERLIAN  
; REGISTRATION NUMBER: 19,683  
; REFERENCE/DOCKET NUMBER: 408.013  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: E (212) 661-8000  
; TELEFAX: (212) 661-8002  
; TELEX:

;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 4617  
;; TYPE: NUCLEIC ACID  
;; STRANDEDNESS: DOUBLE  
;; TOPOLOGY: LINEAR  
;; MOLECULE TYPE: cDNA  
;; ORIGINAL SOURCE:  
;; ORGANISM: RAT  
;; STRAIN:  
;; INDIVIDUAL ISOLATE:  
;; DEVELOPMENTAL STAGE:  
;; HAPLOTYPE:  
;; TISSUE TYPE:  
;; CELL TYPE:  
;; CELL LINE:  
;; ORGANELLE:  
US-08-930-055A-1

Query Match 3.6%; Score 17; DB 4; Length 4617;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 300 AAGACTGACCCAGGCC 316  
Db 3359 AAGACTGACCCAGGCC 3375

RESULT 11  
US-08-930-055A-2  
; Sequence 2, Application US/08930055A  
; Patent No. 6403561  
; GENERAL INFORMATION:  
; APPLICANT: ROSE, CHRISTINANE; VARGAS,  
; APPLICANT: FROYLAN; BOURGAT, PIERRE; SCHWARTZ, JEAN-  
; APPLICANT: CHARLES; BISHOP, PAUL; BAMBAL, RAMESH;  
; APPLICANT: GANELLIN, CHARON; LEBLOND, BERTRAND; MOORE,  
; APPLICANT: ANDREW; CHAN, SUZANNE; ZHAO, LIHUA.  
; TITLE OF INVENTION: TRIPEPTIDYLPEPTIDASE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIERMAN, MUSERLIAN AND LUCAS  
; ADDRESS: LLP  
; STREET: 600 THIRD AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MSDOS  
; SOFTWARE: MICROSOFT WORD 97  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/930,055A  
; FILING DATE: 19-FEB-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR96/00700  
; FILING DATE: 9-MAY-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR95/05489  
; FILING DATE: 9-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CHARLES A. MUSERLIAN  
; REGISTRATION NUMBER: 19,683  
; REFERENCE/DOCKET NUMBER: 408.013  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: E (212) 661-8000  
; TELEFAX: (212) 661-8002  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 5109  
TYPE: NUCLEIC ACID  
STRANDEDNESS: DOUBLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: CDNA  
ORIGINAL SOURCE:  
ORGANISM: RAT  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
US-08-930-055A-2

Query Match 3.6%; Score 17; DB 4; Length 5109;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 300 AGACTGACCCGAGGCC 316  
DB 3851 AAGACTGACCCGAGGCC 3867

## RESULT 12

US-09-313-294A-4713/c  
Sequence 4713, Application US/09313294A  
Patent No. 6476212  
GENERAL INFORMATION:  
APPLICANT: Lalgudi, Raghunath V.  
APPLICANT: Ito, Laura Y.  
APPLICANT: Sherman, Bradley K.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
FILE REFERENCE: PL-0017 US  
CURRENT APPLICATION NUMBER: US/09/313,294A  
CURRENT FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 7600  
SOFTWARE: PERL Program  
SEQ ID NO 4713  
LENGTH: 298  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. 6476212 700348913H1  
NAME/KEY: unsure  
LOCATION: 259, 269, 288  
OTHER INFORMATION: a, t, c, g, or other  
US-09-313-294A-4713

Query Match 3.4%; Score 16; DB 4; Length 298;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 TCCTTGCAGCAGATGT 180  
DB 221 TCTTGCAGCAGATGT 206

## RESULT 13

US-09-621-976-2230  
Sequence 2230, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 2230  
LENGTH: 440  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 32..253  
US-09-621-976-2230

Query Match 3.4%; Score 16; DB 4; Length 440;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 391 CTGCTTTGCTCTCTA 406  
DB 240 CTGCTTTGCTCTCTA 255

## RESULT 14

US-09-134-001C-2724  
Sequence 2724, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 2724  
LENGTH: 813  
TYPE: DNA  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-2724

Query Match 3.4%; Score 16; DB 4; Length 813;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 TCTAATAAGCAATTAT 39  
DB 274 TCTAATAAGCAATTAT 289

## RESULT 15

US-08-956-171E-1111  
Sequence 1111, Application US/08956171E  
Patent No. 6593114  
GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
Gil H. Choi  
Patrick S. Dillon  
Craig A. Rosen  
Steven C. Barash  
Michael R. Fannon  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5256  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

```

;
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 1111:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1111:
US-08-956-171E-1111

```

```

Query Match          3.4%; Score 16; DB 4; Length 825;
Best Local Similarity 100.0%; Fred.No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 418 TGTATTTTCATTGAA 433
   |||||||
Db 265 TGTATTTTCATTGAA 280

```

```

Search completed: July 31, 2004, 10:51:20
Job time : 68 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 08:13:50 ; Search time 65 Seconds  
(without alignments)  
4063.948 Million cell updates/sec

Title: US-10-020-540A-1  
Perfect score: 476  
Sequence: 1 aactatgctgacgaagat.....aacctccctgcttacaaacc 476

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/ptUS COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36.2	7.6	161652	4	US-09-497-855A-40
2	34.6	7.3	7218	1	US-08-232-463-14
3	34.4	7.2	1887	4	US-09-107-532A-2783
4	34	7.1	897	4	US-09-107-532A-72
5	32.4	6.8	1145	3	US-09-227-794-1
6	32.4	6.8	1677	4	US-09-252-991A-4734
7	32.4	6.8	2151	4	US-09-252-991A-4774
8	32	6.7	1923	4	US-09-328-352-3932
9	32	6.7	5170	4	US-09-677-046A-5
10	31.8	6.7	1215	3	US-08-642-274D-35
11	31.8	6.7	1215	3	US-08-952-014C-35
12	31.8	6.7	1482	4	US-09-328-352-2537
13	31.8	6.7	72604	4	US-09-268-992-7
14	31.8	6.7	72604	4	US-09-657-474-7
15	31.6	6.6	191	4	US-09-621-976-18332
16	31.6	6.6	1437	4	US-09-137-223A-4
17	31.6	6.6	16442	3	US-08-781-891-208
18	31.6	6.6	16442	3	US-09-618-166-208
19	31.6	6.6	58407	4	US-08-916-421B-2
20	31.4	6.6	870	4	US-09-107-532A-1977
21	31.4	6.6	2567	3	US-08-993-260-4
22	31.4	6.6	12482	4	US-09-512-563C-25
23	31.4	6.6	50000	4	US-09-145-053-4
24	31.2	6.6	5044	4	US-09-735-935-3
25	31.2	6.6	45716	4	US-08-965-048-5
26	31.2	6.6	45989	4	US-08-965-048-6
27	31	6.5	708	4	US-09-134-001C-1313

C	28	31	6.5	2664	4	US-09-328-352-353	Sequence 353, App
	29	30.8	6.5	201	4	US-09-134-000C-2031	Sequence 2031, App
	30	30.8	6.5	1434	4	US-09-134-000C-304	Sequence 304, App
	31	30.8	6.5	3475	4	US-08-924-829C-15	Sequence 15, Appl
C	32	30.6	6.4	600	4	US-09-621-976-10493	Sequence 10493, A
	33	30.6	6.4	2707	3	US-09-121-964-2	Sequence 2, Appli
	34	30.4	6.4	591	4	US-09-328-352-1319	Sequence 1319, Ap
	35	30.4	6.4	3942	4	US-09-601-198-50	Sequence 50, Appl
C	36	30.4	6.4	4977	1	US-08-030-096-7	Sequence 7, Appli
	37	30.4	6.4	10607	1	US-08-078-090-3	Sequence 3, Appli
C	38	30.4	6.4	202001	4	US-09-734-674-3	Sequence 359, App
	39	30.2	6.3	643	4	US-08-956-171E-359	Sequence 2441, Ap
	40	30.2	6.3	1170	4	US-09-134-000C-2441	Sequence 4, Appli
C	41	30.2	6.3	1737	1	US-08-202-056-4	Sequence 3, Appli
	42	30.2	6.3	1737	1	US-08-076-093A-3	Sequence 3, Appli
C	43	30.2	6.3	1737	1	US-08-701-265-3	Sequence 3, Appli
	44	30.2	6.3	1737	2	US-08-284-586-3	Sequence 3, Appli
C	45	30.2	6.3	1737	2	US-08-805-478-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1  
US-09-497-855A-40  
; Sequence 40, Application US/09497855A  
; Patent No. 6605432  
; GENERAL INFORMATION:  
; APPLICANT: Huang, Tim  
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION  
; FILE REFERENCE: UMO1523  
; CURRENT APPLICATION NUMBER: US/09/497,855A  
; CURRENT FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/120,592  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: 60/118,760  
; PRIOR FILING DATE: 1999-02-05  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 40  
; LENGTH: 161652  
; TYPE: DNA  
; ORGANISM: Homo sapiens;  
US-09-497-855A-40

Query Match	7.6%	Score 36.2;	DB 4;	Length 161652;
Best Local Similarity	46.9%;	Pred. No. 2.7;		
Matches 113;	Conservative	0;	Mismatches 128;	Indels 0;
Gaps 0;				
QY	152	CACCAAGCCACATCTTTGCAGCAGATGTGAAGCAGCCAGAGTGTCCACAGACGCACT	211	
Db	21744	CACTCTCCCTTCTTATTCACATAGGACTTAAGTCTTAGCATTTAGTAAAGAAAGA	21803	
QY	212	CAGAAAGGCACTTTTACCGACACAGAAAAGAACACACAGCTCATCTCAACATGT	271	
Db	21804	AACAAAGGCATCTTAATCAGAAAAGGGAAGTAAATGATCTCTGTTTTAGATGACTT	21863	
QY	272	AGACTCTCTTATCGCTCGGCTCAGATTAAGTGTACCCAGCCAGCACTAAAGAGAA	331	
Db	21864	GATCTTACATGAGAAAATCCCTAAGACACCCACAAAATTTGTTAGAACTTAATTA	21923	
QY	332	ATAATCAAGTGTCTTAGCTCCACTTTAGCTTTTAAATTAATTTGTTTCAATTTATTC	391	
Db	21924	ATTCACTCAAGTGTGCAGATACAAATCAACATATAAAATCAGTTGTTATTTTGTACAC	21983	
QY	392	T 392		
Db	21984	T 21984		

RESULT 2  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463

Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZgpt-F1s  
US-08-232-463-14

Query Match 7.3%; Score 34.6; DB 1; Length 7218;  
Best Local Similarity 2.0%; Pred. No. 1.9;  
Matches 7; Conservative 192; Mismatches 146; Indels 0; Gaps 0;

QY 10 GATGACAAGATAATCTCTAATAGCAATATTTCAGAAATTAATCAAGGAGAAAGAAATTAATA 69  
DB 1399 RRR 1340  
QY 70 ACTCTTTTCAGAAATATGAAGCCCGCTTTTACAGAGTGGCCAGTAGCTATCATCTGAAAGACA 129  
DB 1339 RR 1280  
QY 130 GCAAGACAATGGTGTCTGATGACCAAGACACATCTTTGCACGACATGTGAGCAGCC 189  
DB 1279 RR 1220  
QY 190 AGAGTGGTCCACAGACCACTCAAGAAAGCATCTTTTACCGACACAGAAAAGACAAC 249  
DB 1219 RR 1160  
QY 250 CACAGCTCATCATCAACATCTAGACTCTCGTTATGCTCGGTGGAAGATAAGACTGACC 309  
DB 1159 RR 1100  
QY 310 CCAGCCGACGACTAAGAGAAATAATGCAAGTGGTCTAGCTCC 354  
DB 1099 RR 1055

RESULT 3  
US-09-107-532A-2783/c  
; Sequence 2783, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Denek  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 2783:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1887 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...1887  
SEQUENCE DESCRIPTION: SEQ ID NO: 2783:  
US-09-107-532A-2783

Query Match 7.2%; Score 34.4; DB 4; Length 1887;  
Best Local Similarity 57.4%; Pred. No. 1.1;  
Matches 62; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 323 AAAGAAGAAATTAATGCAAGTGGTCTAGCTCCACTTTAGCTTTAATAATATGTTTCATT 382  
DB 511 AAACAAAATAAAATAAAGTAGTCGTAATTTAGCTCATCTCTCTATTATTCTTAGTCCTT 452  
QY 383 ATTATTCTCTGCTTTTGTCTCTATATAAAGAGCTTGTTATTTTCATT 430  
DB 451 TATTGGTGTGCTGTATCTTTTATCATCAAAATCATTCATCTCTT 404

RESULT 4  
US-09-107-532A-72  
; Sequence 72, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:

```
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 897 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...897
; SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-09-107-532A-72
Query Match 7.1%; Score 34; DB 4; Length 897;
Best Local Similarity 52.1%; Pred. No. 1.1; Indels 0; Gaps 0;
Matches 76; Conservative 0; Mismatches 70;
QY 1 AACTATGCTGATGACAAGATAATCTTAATAGCAATTTATTGAGAAATTAATCAAGAGAAA 60
Db 268 AAAGAGGTGGCTAAACGGTTAAATGTCAAATCTCTTTAAAGAATAACACGTGATGGA 327
QY 61 GAATTAATACTCTTTCAGATATAGAGCCGCTTTACAGTGGCCAGCTAGCTATCACT. 120
Db 328 CAATATCACTTCTGTTTCAAAAGAGAAATCAGATTTTGCTGCAACGATTTTGGTATCACA 387
QY 121 GAAAGACAGCAAGACAATGGTGCT 146
Db 388 AAAGAAAGAGAAAAAATTCCTCT 413
RESULT 5
US-09-227-794-1/c
; Sequence 1, Application US/09227794A
; Patent No. 617613
; GENERAL INFORMATION:
; APPLICANT: Coughlan, Sean J.
; APPLICANT: Winfrey, Ronnie J. Jr.
; TITLE OF INVENTION: Seed-Preferred Promoter
```

```
; FILE REFERENCE: 0945
; CURRENT APPLICATION NUMBER: US/09/227,794A
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: US 09/227,794
; EARLIER FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1145
; TYPE: DNA
; ORGANISM: Glycine max
US-09-227-794-1
Query Match 6.8%; Score 32.4; DB 3; Length 1145;
Best Local Similarity 52.2%; Pred. No. 3.6;
Matches 72; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 296 AGATAAGACTGACCCCGCCAGCAGCACTAAAGAGAATAATGCAAGTGGTCTTAGCTCCA 355
Db 459 AGATCAATTGAGTAGAATCATCTTCTGTAAACAATAATGAAGTGGTTAAAGCATAT 400
QY 356 CTTTAGCTTTAATAATTATGTTTTCATTATTATCTCTGCTTTTCTCTCTATATAAGAG 415
Db 399 TTTTGTATTTTACTTTTTTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 340
QY 416 CTTGTATTTTTCATTGAA 433
Db 339 AAGTCTCTTCTTTTGTGA 322
RESULT 6
US-09-252-991A-4734/c
; Sequence 4734, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4734
; LENGTH: 1677
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4734
Query Match 6.8%; Score 32.4; DB 4; Length 1677;
Best Local Similarity 52.2%; Pred. No. 4.4;
Matches 72; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 216 AAAGGATCTTTCTTACCGACACAGAAAAAGACAAACACAGCTCATCATCAACATGTAGAC 275
Db 1539 AAGGGATGACCTACCGGAGGAGAAATTCCTGGGAGCTCAAGGACCTGGCCGAAGAG 1480
QY 276 TGTCTGTATGCTGCGCTGAAGATAAGACTGACCCGAGCCAGCAGCACTAAAGAGAAATAA 335
Db 1479 TGGCGTTCCAGCATGCTGCGAGGCTCGGCCGGAAGCCAGCAAGAGCTGATGAACAGTAC 1420
QY 336 TGCAAGTGGTCTTAGCTC 353
Db 1419 CTGGAAGGGCGAGCTC 1402
RESULT 7
US-09-252-991A-4774
; Sequence 4774, Application US/09252991A
; Patent No. 6551795
```



APPLICANT: Shiloh, Yosef  
TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE AND ITS  
TITLE OF INVENTION: GENOMIC ORGANIZATION  
NUMBER OF SEQUENCES: 91  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kohn & Associates  
STREET: 30500 No. 6265158thwestern Hwy., Suite 410  
CITY: Farmington Hills  
STATE: Michigan  
COUNTRY: U.S.  
ZIP: 48334  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/952,014C  
APPLICATION NUMBER: US/08/952,014C  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kohn, Kenneth I.  
REGISTRATION NUMBER: 30,995  
REFERENCE/DOCKET NUMBER: 2290.00028  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 810-539-5050  
TELEFAX: 810-539-5055  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1215 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-952-014C-35

Query Match 6.7%; Score 31.8; DB 3; Length 1215;  
Best Local Similarity 50.3%; Pred. No. 5.7; Indels 0; Gaps 0;  
Matches 78; Conservative 0; Mismatches 77; Indels 0; Gaps 0;  
QY 320 ACTAAGAAGAAATAATGCAAGTGGCTCTAGCTCCACTTTAGCTTTAATAATTATGTTTC 379  
DB 277 AGTTAAAGCAATTTTGAATAATCTTCCAAAAGCCCTGTAAGTATACATGATGAGTTTA 336  
QY 380 ATTATTAATCTGCTTTTCTCTCTATATAAGAGCTTCTATTTCAATTTGAGGCAGA 439  
DB 337 ATAATAGAACATTCCTCTCTTTTGTAGCTAAAGAACTTTGTAAATACATCTTAAAGAGGA 396  
QY 440 GGCGAACACACACACAGACCTCCCTGCTTACAAA 474  
DB 397 AAGTAACAAGAAAGAAATTTATCTCATATTA 431

RESULT 12  
US-09-328-352-2537  
Sequence 2537, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 2537  
LENGTH: 1482  
TYPE: DNA  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-2537

Query Match 6.7%; Score 31.8; DB 4; Length 1482;  
Best Local Similarity 52.7%; Pred. No. 6.3;

Matches 69; Conservative 0; Mismatches 62; Indels 0; Gaps 0;  
QY 3 CTATGCTGATCACAAGATTAATCTTAATAAGCAATTTATTCAGATTAATCAAGGAGAAAGA 62  
DB 120 CTTTACCATGAAAAAATTTGTAATTTACTTTAAGATCAACATGATCTGAATTTGGTAAC 179  
QY 63 ATTAATACTCTTTTCAGAAATATGAAGCCCGCTTTTACAAGTGGCCAGCTAGCTATCACTGA 122  
DB 180 AGTCAAAACATTTAAAGTTTATTATACCGATATGCAAAAGTGAATCCAGCTATGTCTAA 239  
QY 123 AAAGACAGCAA 133  
DB 240 AAATACTGCTA 250  
RESULT 13  
US-09-268-992-7/c  
Sequence 7, Application US/09268992  
Patent No. 6342351  
GENERAL INFORMATION:  
APPLICANT: Chen, H.  
APPLICANT: Freimer, N.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING  
TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS  
FILE REFERENCE: 7853-138  
CURRENT APPLICATION NUMBER: US/09/268,992  
CURRENT FILING DATE: 1999-03-16  
EARLIER APPLICATION NUMBER: 09/236,134  
EARLIER FILING DATE: 1999-01-22  
EARLIER APPLICATION NUMBER: 60/106,056  
EARLIER FILING DATE: 1998-10-28  
EARLIER APPLICATION NUMBER: 60/088,312  
EARLIER FILING DATE: 1998-06-05  
EARLIER APPLICATION NUMBER: 60/078,044  
EARLIER FILING DATE: 1998-03-16  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 7  
LENGTH: 72604  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: modified base  
LOCATION: all n positions  
OTHER INFORMATION: n=a, c, g, or t  
US-09-268-992-7

Query Match 6.7%; Score 31.8; DB 4; Length 72604;  
Best Local Similarity 49.7%; Pred. No. 40;  
Matches 81; Conservative 0; Mismatches 82; Indels 0; Gaps 0;  
QY 278 TCGTTATGCGTGGCTGAAGATAGACTGACCCAGGCCAGCACTAAGAGAAATATG 337  
DB 23019 TAGTTCTTCCTGGTCATAGATAAGAACACAGACCTAACTGAGCTAAGGACCAAAAAAT 22960  
QY 338 CAAGTGGCTCTAGCTCCACTTTAGCTTTAATAATTTATTTCAATTTATTTCTCTGCTTT 397  
DB 22959 CTGCATCACTGTTGCAAGTTTACTTTAGATTTTAACTTATAGTTTATGTTAAATTAC 22900  
QY 398 TCGCTCTATATAAGAGCTTGTATTTCATTTGAGGCAGAG 440  
DB 22899 TGTTCATTTTGAAGAGTACTATTTCGTTATTATAAAGTG 22857

RESULT 14  
US-09-657-474-7/c  
Sequence 7, Application US/09657474  
Patent No. 6399762  
GENERAL INFORMATION:  
APPLICANT: Chen, H.  
APPLICANT: Freimer, N.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING  
TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 08:12:51 ; Search time 335 Seconds  
(without alignments)  
6966.850 Million cell updates/sec

Title: US-10-020-540A-1

Perfect score: 476

Sequence: 1 aactatgctgatacaagat.....aaactccctgcttacaacc 476

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	476	100.0	476	14	US-10-020-540A-1
2	476	100.0	518	14	US-10-020-540A-2
3	369	77.5	651	14	US-10-020-540A-3
4	369	77.5	701	14	US-10-020-540A-4
5	40.2	8.4	539	13	US-10-027-632-72822
6	40.2	8.4	539	13	US-10-027-632-72821
7	40.2	8.4	539	16	US-10-027-632-72822
8	40.2	8.4	539	16	US-10-027-632-312801
9	40.2	8.4	547	13	US-10-027-632-49899
10	40.2	8.4	547	16	US-10-027-632-49899
11	37	7.8	90798	17	US-10-318-819A-4
12	37	7.8	3673778	15	US-10-312-841-2
13	36.8	7.7	6012	13	US-10-221-714A-449
14	36.6	7.7	17211	15	US-10-311-455-628

C	15	36.4	7.6	1141	13	US-10-424-599-93424	Sequence 93424, A
	16	36.2	7.6	161652	15	US-10-081-327-40	Sequence 40, Appl
	17	35.8	7.5	379	13	US-10-424-599-132308	Sequence 132308, A
	18	35.8	7.5	422	9	US-09-783-590-4784	Sequence 4784, Ap
C	19	35.6	7.5	586	13	US-10-027-632-105291	Sequence 105291, A
C	20	35.6	7.5	586	16	US-10-027-632-105291	Sequence 105291, A
	21	35.4	7.4	593	13	US-10-027-632-214128	Sequence 214128, A
	22	35.4	7.4	593	16	US-10-027-632-214128	Sequence 214128, A
	23	35	7.4	458	9	US-09-764-853-272	Sequence 272, App
	24	35	7.4	458	9	US-09-764-860-86	Sequence 86, Appl
	25	35	7.4	458	15	US-10-091-438-53	Sequence 53, Appl
	26	35	7.4	458	15	US-10-074-095-86	Sequence 86, Appl
	27	35	7.4	458	16	US-10-212-872-86	Sequence 86, Appl
	28	35	7.4	494	9	US-09-783-590-4851	Sequence 4851, Ap
	29	35	7.4	502	9	US-09-764-853-15	Sequence 15, Appl
	30	35	7.4	504	9	US-09-764-860-770	Sequence 770, App
	31	35	7.4	504	15	US-10-091-438-265	Sequence 265, App
	32	35	7.4	504	15	US-10-074-095-770	Sequence 770, App
	33	35	7.4	504	16	US-10-212-872-770	Sequence 770, App
	34	35	7.4	573	13	US-10-027-632-270337	Sequence 270337, A
	35	35	7.4	729	13	US-10-027-632-15321	Sequence 15321, A
C	36	35	7.4	729	16	US-10-027-632-15321	Sequence 15321, A
C	37	35	7.4	1344	13	US-10-425-114-6401	Sequence 6401, Ap
	38	35	7.4	2693	13	US-10-424-599-97948	Sequence 97948, A
C	39	35	7.4	655	16	US-10-341-961A-245	Sequence 245, App
C	40	34.8	7.3	6314	9	US-09-764-887-473	Sequence 473, App
C	41	34.6	7.3	6314	15	US-10-073-961-473	Sequence 473, App
C	42	34.6	7.3	9289	9	US-09-764-887-471	Sequence 471, App
C	43	34.6	7.3	9289	15	US-10-073-961-471	Sequence 471, App
C	44	34.6	7.3	17865	9	US-09-764-887-472	Sequence 472, App
C	45	34.6	7.3	17865	15	US-09-764-887-472	Sequence 472, App

ALIGNMENTS

RESULT 1  
US-10-020-540A-1  
; Sequence 1, Application US/10020540A  
; Publication No. US20020182593A1  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Gusui  
; TITLE OF INVENTION: Strawberry Vein Banding Virus Promoter  
; TITLE OF INVENTION: (SVBV) Promoter  
; FILE REFERENCE: 0173.210us  
; CURRENT APPLICATION NUMBER: US/10/020,540A  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/245,354  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 476  
; TYPE: DNA  
; ORGANISM: Strawberry Vein Banding Virus (SVBV) Strain E3  
US-10-020-540A-1

Query Match	100.0%;	Score 476;	DB 14;	Length 476;
Best Local Similarity	100.0%;	Pred. No. 1e-127;		
Matches	476;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	AACATGCTGATGACAAAGATAATCTTATAGCANTTATTCAGATTATTCAGAGGAAA	60	
DB	1	AACATGCTGATGACAAAGATAATCTTATAGCANTTATTCAGATTATTCAGAGGAAA	60	
QY	61	GAATTAATAACCTCTTTCAGATAATCAAGCCGCTTTACAAGTGGCCAGCTAGCTACT	120	
DB	61	GAATTAATAACCTCTTTCAGATAATCAAGCCGCTTTACAAGTGGCCAGCTAGCTACT	120	
QY	121	GAAGAGACAGCAAGCAATGGTGTTCGATGACAGAACACATCTTTGACAGAGATGT	180	
DB	121	GAAGAGACAGCAAGCAATGGTGTTCGATGACAGAACACATCTTTGACAGAGATGT	180	







```
US-10-027-632-72822/c
; Sequence 72822, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72822
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-72822
```

```
Query Match      8.4%; Score 40.2; DB 16; Length 539;
Best Local Similarity 53.5%; Pred. No. 0.41;
Matches 84; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 233 ACACAGAAAAGACACACAGCTCATCCACATGTAGACTGCTGTTATGCGTCGGC 292
Db 423 AAACAAAAAAGAAAACCTCAGGCCAATATCCATGTGACATTCGACACAAAATCTCA 364
QY 293 TGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAAATAATGCAAGTGGTCTAGCT 352
Db 363 TTAATAATAAGTCAACCAATCCAGCAGCAATCAGAAAGTTTATCTACCATATCAAG 304
QY 353 CCACCTTTAGCTTTAATAATTATGTTTCATTATTTC 389
Db 303 TTGGCTTCACTCCAGGCATTTTATTCAGTTTTTTAC 267
```

```
RESULT 8
US-10-027-632-312801/c
; Sequence 312801, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
```

```
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 312801
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-312801
```

```
Query Match      8.4%; Score 40.2; DB 16; Length 539;
Best Local Similarity 53.5%; Pred. No. 0.41;
Matches 84; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 233 ACACAGAAAAGACACACAGCTCATCCACATGTAGACTGCTGTTATGCGTCGGC 292
Db 423 AAACAAAAAAGAAAACCTCAGGCCAATATCCATGTGACATTCGACACAAAATCTCA 364
QY 293 TGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAAATAATGCAAGTGGTCTAGCT 352
Db 363 TTAATAATAAGTCAACCAATCCAGCAGCAATCAGAAAGTTTATCTACCATATCAAG 304
QY 353 CCACCTTTAGCTTTAATAATTATGTTTCATTATTTC 389
Db 303 TTGGCTTCACTCCAGGCATTTTATTCAGTTTTTTAC 267
```

```
RESULT 9
US-10-027-632-49899/c
; Sequence 49899, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49899
; LENGTH: 547
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-49899
```

```
Query Match      8.4%; Score 40.2; DB 13; Length 547;
Best Local Similarity 53.5%; Pred. No. 0.41;
Matches 84; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 233 ACACAGAAAAGACACACAGCTCATCCACATGTAGACTGCTGTTATGCGTCGGC 292
Db 431 AAACAAAAAAGAAAACCTCAGGCCAATATCCATGTGACATTCGACACAAAATCTCA 372
QY 293 TGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAAATAATGCAAGTGGTCTAGCT 352
Db 371 TTAATAATAAGTCAACCAATCCAGCAGCAATCAGAAAGTTTATCTACCATATCAAG 312
QY 353 CCACCTTTAGCTTTAATAATTATGTTTCATTATTTC 389
```

Db 311 TTGGCTTCACTCCAGGCAATTTATTTCAAGTTTATTTTAC 275

## RESULT 10

US-10-027-632-49899/c  
; Sequence 49899, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; Polymorphisms in the Human Genome  
; FILE REFERENCE: 108927.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 49899  
; LENGTH: 547  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-49899

Query Match 8.4%; Score 40.2; DB 16; Length 547;  
Best Local Similarity 53.5%; Pred. No. 0.41;  
Matches 84; Conservative 0; Mismatches 73; Indels 0; Gaps 0;  
  
QY 233 ACACGAGAAAGACACACAGCTCATCCACATGTAGATGTCGTTATCGCGCGC 292  
Db 431 AACAAAAAGAAACCTCAGGCCAATATCCATGATGACATGACAAAAATCCTCA 372  
QY 293 TGAAGATAAGACTGACCCAGGCGCAGCATAAAGAGAAATATGCAAGTGTCTTAGCT 352  
Db 371 TTAATAATAGTCAACCAATCCAGCAGCAATCAGAAAGTTATCTACCATATCAAG 312  
QY 353 CCACCTTAGCTTAAATATATGTTTCAATTTATTC 389  
Db 311 TTGGCTTCACTCCAGGCAATTTATTTCAAGTTTATTTTAC 275

## RESULT 11

US-10-318-819A-4  
; Sequence 4, Application US/10318819A  
; Publication No. US20040115645A1  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Kenneth W. Dobie  
; TITLE OF INVENTION: MODULATION OF DRK2 EXPRESSION  
; FILE REFERENCE: PFS-0069  
; CURRENT APPLICATION NUMBER: US/10/318,819A  
; CURRENT FILING DATE: 2002-12-12  
; NUMBER OF SEQ ID NOS: 133  
; SEQ ID NO 4  
; LENGTH: 90798  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1349, 1350, 13454-13553, 28775-28874, 46851-46950, 52633-52732, 67991-  
; LOCATION: 68090, 85646-85745, 87254-87353, 89363-89462, 90660-90759

; OTHER INFORMATION: n = A, T, C or G  
US-10-318-819A-4

Query Match 7.8%; Score 37; DB 17; Length 90798;  
Best Local Similarity 51.5%; Pred. No. 60;  
Matches 85; Conservative 0; Mismatches 80; Indels 0; Gaps 0;  
  
QY 235 ACAGAAAAAGACACACAGCTCATCCACATGTAGATGTCGTTATCGCGCGTG 294  
Db 52993 AACAAAAAGACCAATTTATTCCTAAAGGTAAGCTAACATTTTCATGCTT 53052  
QY 295 AAGATAAGACTGACCCAGGCGCAGCACTAAAGAGAAATAATGCAAGTGTCTTAGCTCC 354  
Db 53053 AAAATAGATGCTAAATCACCTTTTAATAAAAAATTACATTTCAAGTAATCCTAAGTAT 53112  
QY 355 ACTTAGCTTTAAATATATGTTTCATTTATTTATTTCTGCTTTTG 399  
Db 53113 AATTTTTTTTTTACTATCATCAGGAGTTAAACTACTCTCTCTTTG 53157

## RESULT 12

US-10-312-841-2/c  
; Sequence 2, Application US/10312841  
; Publication No. US20030186277A1  
; GENERAL INFORMATION:  
; APPLICANT: Epigenomics AG  
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC  
; FILE REFERENCE: E01/1208/WO  
; CURRENT APPLICATION NUMBER: US/10/312,841  
; CURRENT FILING DATE: 2002-12-30  
; NUMBER OF SEQ ID NOS: 2  
; SEQ ID NO 2  
; LENGTH: 3673778  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (379615)  
US-10-312-841-2

Query Match 7.8%; Score 37; DB 15; Length 3673778;  
Best Local Similarity 48.0%; Pred. No. 3.9e+02;  
Matches 106; Conservative 0; Mismatches 115; Indels 0; Gaps 0;  
  
QY 48 AATCAAGAGAGAAAGAAATTAATACTCTTTCAAGATATGAAGCCGCTTTACAGTGGCCA 107  
Db 3196228 AATAAACCATTAATAATAATTAATCTTAACACCATACCCTCCCATATAATTCATAACA 3196169  
QY 108 GCTAGCTATCACTGAAAGACAGCAAGACAATGTGTCTCGATGACCCAGAACCAATCT 167  
Db 3196168 CCTATCAAAAAACGAAAAACATAAACAATAAATAAATACGCCATAAAAAAAACT 3196109  
QY 168 TTGCAGCAGATGTGAAGCAGCCAGAGTGTCCACAAGACGCACTCAGAAAAAGGATCTTC 227  
Db 3196108 CTTTACCTATTTAAAAATCTTCGTATCTCTAATAAAAAAATAAAAAACACAATATC 3196049  
QY 228 TACCGACACAGAAAAAGACACACAGCTCATATCCAAACA 268  
Db 3196048 AATATAAATCAATATAAATTTCAATATCATCTACACCCCAACA 3196008

## RESULT 13

US-10-221-714A-449/c  
; Sequence 449, Application US/10221714A  
; Publication No. US20040048254A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PFEENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with  
; tumor suppressor genes and oncogenes

```
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221,714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 449
; LENGTH: 6012
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-449

Query Match
Best Local Similarity 7.7%; Score 36.8; DB 13; Length 6012;
Matches 110; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 34 AATTATTCAGATTAATCAAGGAGAGAAATTAATTAATCTTTTCAAGATATGAAGCCCGC 93
Db 861 ACTCATTTTACAAAAAACAACAAACAACTAAATTAATCTTACCCCAAAATACAAAAACACC 802
QY 94 TTTCAGTGGCCAGCTAGCTATCACTGAAAGACAGACAAATGGTGTCTCGATGCA 153
Db 801 AAAACCACTACCAAAACCTATACCCAAATACAAAAAATAATCTCAAAAA 742
QY 154 CAGAACCAATCTTTGCGAGATGTGAAGACGCGAGTGTGTCACAGACGCACTCA 213
Db 741 CAAAAAATATACGACAAAAAACAACAAAAAACCCTACATAAAAAAACCACCA 682
QY 214 GAAAGGATCTTCTACCGACACAGAAAAACACACACACCTCATCTCA 265
Db 681 ATAAAAACACACACACACACCCCAATTAACACACACACACCCCAATAAACA 630

RESULT 14
US-10-311-455-628/c
; Sequence 628, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 628
; LENGTH: 17211
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-628

Query Match
Best Local Similarity 7.6%; Score 36.4; DB 13; Length 1141;
Matches 79; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 313 GGCCAGCAGCTAAAGAGAAATTAATCAAGTGTCTCTAGCTCCACTTTTACCTTTTAAATTT 372
Db 1141 GGCCGCCCTTTTCAAGCAGAAAAAGACATTTATAACCAAGTCTCTGCAAACTCACAAATTT 1082
QY 373 ATGTTTCATTATTCTCTGCTTTTGTCTCTCTATATAAGAGCTTGTATTTCATTGA 432
Db 1081 CAACTCCGGTTAACTTCATTTGCTTTTACGTATAAGAGATTGCAATTTCTTTTGA 1022
QY 433 AGGCAGAGCGGAAACACACACAGAACCTC 462
Db 1021 AAACCTAGATCAACAAAAAATTTTC 992

Search completed: July 31, 2004, 09:31:32
Job time : 340 secs

Query Match
Best Local Similarity 56.1%; Pred. No. 31;
Matches 69; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 17 AGATAATTCCTAATAAGCAATTAATTCAGATTAATCAAGGAGAGAAAGAAATTAATAACTCTTT 76
Db 7211 AATAATCTCCACCTCATCTTCGAAATAACTAATAATTAATAATTTTCAAAACTTT 7152
QY 77 CAGAATATGAAGCCCGCTTTTCAAGTGGCCAGCTAGCTATCACTGAAAGACAGCAAGAC 136
Db 7151 TATATAATAATATCTCTATATAATTTCCCAACTCTCTATAACTATAAATTTCTCTATAA 7092
QY 137 AAT 139
Db 7091 AAT 7089

RESULT 15
US-10-424-599-93424/c
; Sequence 93424, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with Diagnosis of Plant Diseases
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 93424
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_55376C.1
US-10-424-599-93424
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 07:54:50 ; Search time 1998 Seconds  
(without alignments)  
7114.313 Million.cell updates/sec

Title: US-10-020-540A-1  
Perfect score: 476  
Sequence: 1 aactatgctgatacaagat.....aacctccctgcttacaacc 476

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

- 1: em\_estba.\*
- 2: em\_esthum.\*
- 3: em\_estlin.\*
- 4: em\_estm.\*
- 5: em\_estov.\*
- 6: em\_estpl.\*
- 7: em\_estro.\*
- 8: em\_estc.\*
- 9: gb\_estl.\*
- 10: gb\_est2.\*
- 11: gb\_est3.\*
- 12: gb\_est4.\*
- 13: gb\_est5.\*
- 14: gb\_est5.\*
- 15: em\_estfun.\*
- 16: em\_estcom.\*
- 17: em\_gss\_hum.\*
- 18: em\_gss\_inv.\*
- 19: em\_gss\_pln.\*
- 20: em\_gss\_vrt.\*
- 21: em\_gss\_fun.\*
- 22: em\_gss\_man.\*
- 23: em\_gss\_mus.\*
- 24: em\_gss\_pro.\*
- 25: em\_gss\_rod.\*
- 26: em\_gss\_phg.\*
- 27: em\_gss\_vrl.\*
- 28: gb\_gss1.\*
- 29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	54.8	11.5	994	13	BX414650
C 2	50.6	10.6	712	13	BX416727
C 3	44.4	9.3	886	29	CNS017XV
C 4	42.2	8.9	763	14	CA228227

5	41.4	8.7	1201	13	BX356232
C 6	40.8	8.6	919	29	CNS006S5
C 7	40.4	8.5	983	13	BX432706
C 8	40.4	8.5	1044	13	BX336333
C 9	40.2	8.4	648	13	BX417150
C 10	40.2	8.4	852	13	BX393687
C 11	40	8.4	1201	13	BX335650
C 12	39.8	8.4	775	29	CC484903
C 13	39.6	8.3	700	29	CE603042
C 14	39.6	8.3	1201	9	AL544517
C 15	39.2	8.2	885	13	BX425603
C 16	38.8	8.2	549	29	AG223026
C 17	38.8	8.2	870	28	AZ693178
C 18	38.8	8.2	1101	29	CNS00KX2
C 19	38.6	8.1	1101	29	CNS008SU
C 20	38.4	8.1	187	12	BG981764
C 21	38.2	8.0	682	29	CC820164
C 22	38.2	8.0	807	14	CF447957
C 23	38.2	8.0	834	29	CNS0102S
C 24	38.2	8.0	958	10	BE777969
C 25	38	8.0	355	28	AQ044946
C 26	38	8.0	1059	13	BX359119
C 27	37.8	7.9	181	12	BI042556
C 28	37.6	7.9	826	29	CC572843
C 29	37.4	7.9	513	29	TA78C11P
C 30	37.4	7.9	593	29	CPA562119
C 31	37	7.8	527	29	CB497353
C 32	36.8	7.7	288	12	BJ392146
C 33	36.8	7.7	679	29	CE147815
C 34	36.8	7.7	688	28	B2452924
C 35	36.8	7.7	1201	13	BX461813
C 36	36.6	7.7	470	12	BU429271
C 37	36.6	7.7	477	12	BI324001
C 38	36.6	7.7	499	10	AW332217
C 39	36.6	7.7	511	10	AW335137
C 40	36.6	7.7	516	10	AW334251
C 41	36.6	7.7	522	10	AW335222
C 42	36.6	7.7	525	10	AW332905
C 43	36.6	7.7	534	10	AW334379
C 44	36.6	7.7	534	12	BU433770
C 45	36.6	7.7	550	10	AW332253

ALIGNMENTS

RESULT 1 BX414650 994 bp mRNA linear EST 15-MAY-2003  
LOCUS BX414650 Homo sapiens THYMUS cDNA clone CSOCAP001YN02  
DEFINITION 3-PRIME mRNA sequence.  
ACCESSION BX414650  
VERSION BX414650.1 GI:30763455  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 994)  
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 6015.f  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/InvitroGen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CSOCAP001DG01NP1.  
Location/Qualifiers  
1. .994



```

/clone_lib="Saccharum officinarum FL3"
/lab_host="DH10B"
/clone="SCGFJ303C03"
/clone_size="10000"
/note="Organ: Base of developing inflorescence (5cm-long) ;
vector: pSport1; Site_1: SalI; Site_2: NotI; An
unidirectional cDNA library generated from [Base of
developing inflorescence (5cm-long)]. cDNA was prepared

```







```
QY 401 TCTCTATATAGAGCTGTGTTTTCATTTGAAGCAGAGCGGACACAC 450
Db 487 TWTHHHHHHMCVTTTTTTTTTTTTTTTTTTTGGACAAGCTTAAACAC 438

RESULT 8
BX333633/c
LOCUS
DEFINITION BX333633 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DC028Y011 3-PRIME, mRNA sequence.
ACCESSION BX333633
VERSION BX333633
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1044)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10303.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DC028AH06NP1&cluster=10303.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DC028AH06NP1.

FEATURES
source
1..1044
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC028Y011"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 8.5%; Score 40.4; DB 13; Length 1044;
Best Local Similarity 30.3%; Pred. No. 51;
Matches 72; Conservative 57; Mismatches 109; Indels 0; Gaps 0;

QY 5 ATGCTGATGACAGATAATTCTTAAGCAATTATTCAGAAATTAATCAAGGAGAAAT 64
Db 591 ATTHKAAAAAASACTTCTCCCAAGSKAASAKKAKSKAKSKAKMKXGMWAGAAA 532

QY 65 TAATAACTCTTTCAGATATGACCCGCTTACAGTGGCCAGCTAGCTATCACTGAA 124
Db 531 ACAMAAAKKAKKASSASATAAATACTCTAMVKKCCCTTCSAASKMSAGSKMKKVASAM 472

QY 125 AGACAGCAAGACAATGTTGTCGATGCACAGAACCAACATCTTTGACGAGAGATGTGAAG 184
Db 471 TAAAGSGAGGGAAMAKMTATCTKGTAAASSAMAAATTTAATTASCMAKMSAKATWAKS 412

QY 185 CAGCAGAGTGGTCCACAGAGCGCACTCAGAAAGGCGATCTTACCGACACAGAAA 242
Db 411 AAMSSCNVGGGTWGMMAACSGSGSGCCCGAMAGCMACMMASSASACAMSGGAGAAA 354

RESULT 9
BX417150/c
LOCUS
DEFINITION BX417150 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE005YMO1
5-PRIME, mRNA sequence.

QY 401 TCTCTATATAGAGCTGTGTTTTCATTTGAAGCAGAGCGGACACAC 450
Db 487 TWTHHHHHHMCVTTTTTTTTTTTTTTTTTTTGGACAAGCTTAAACAC 438
```

```
ACCESSION BX417150
VERSION BX417150.1 GI:30658341
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 648)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE005AG01QPI.

FEATURES
source
1..648
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE005YMO1"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoR V sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 8.4%; Score 40.2; DB 13; Length 648;
Best Local Similarity 23.9%; Pred. No. 60;
Matches 74; Conservative 72; Mismatches 163; Indels 0; Gaps 0;

QY 122 AAAAGACAGACAGACAATGTTGTCGATGCACAGAACCAACATCTTTGCACGACATG 181
Db 588 AAAAGAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGA 529

QY 182 AAGCAGCCAGAGTGTCCACAGAGCGCACTCAGAAAAGGCATCTTCTACCGACACAGAAA 241
Db 528 MMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMM 469

QY 242 AAGAACACACAGCTCATCCACATGACACTGCTGTTATGCTGCGCTGAAGATAA 301
Db 468 MMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMM 409

QY 302 GACTGACCCCGCCAGCAGCACTAAGAGAGAAATAATGCAAGTGTCTCTAGCTCCACTTTAG 361
Db 408 MMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMM 349

QY 362 CTTTAATAATGTTTTCATTTATTTATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 421
Db 348 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 289

QY 422 TTTTCATTT 430
Db 288 HTTYTTTTT 280

RESULT 10
BX393687/c
LOCUS
DEFINITION BX393687 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DC001YG07 5-PRIME, mRNA sequence.
ACCESSION BX393687
VERSION BX393687.1 GI:30624032
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```



Clones are derived from the bovine BAC library CHORI-240 (http://www.chori.org/bacpac/bovine240.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/ordering/information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the British Columbia Genome Sciences Centre, Canada.

Plate: 314 row: I column: 5

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1. .775

/organism="Bos taurus"

/mol\_type="genomic DNA"

/strain="Breed: Hereford"

/db\_xref="taxon:9913"

/clone="CH240\_31415"

/sex="Male"

/cell\_type="Blood"

/clone\_lib="CHORI-240"

/notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;

Hereford bull Li Domino 99375; CHORI-240 Bovine BAC

library (Male) produced by Pieter de Jong"

## ORIGIN

Query Match 8.4%; Score 39.8; DB 29; Length 775;  
Best Local Similarity 60.7%; Pred. No. 73;  
Matches 65; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 333 TAATGCAAGTGGTCCAGTCCACTTAGCTTTAATAATATGTTTCATTAATTTCT 392

Db 385 TCATACAGACATCTAGTTCATGATATATTAATTAATTTCTTTTAGGTTGT 326

QY 393 GCATTTCCTCTATATAAGAGTGTGTTATTTTCATTGAGGCACA 439

Db 325 GCATTTCATGCTGTAGAAATTCCTTCCTTGGAGTCATA 279

## RESULT 13

CE603042/c

LOCUS

DEFINITION tigr-gss-dog-17000356692102 Dog Library Canis familiaris genomic,  
genomic survey sequence.

ACCESSION CE603042

VERSION CE603042.1 GI:36919881

SOURCE GSS.

ORGANISM Canis familiaris (dog)

Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

1 (bases 1 to 700)

Ruschn,D.B., Deicher,A.L., Pop,M., Wang,W., Fraser,C.M. and

Venter,J.C.

The dog genome: survey sequencing and comparative analysis

Science 301 (5641), 1998-1903 (2003)

22875432

PUBLISHED 14512627

Contact: Kirkness EF

The Institute for Genomic Research

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,

Rockville, MD 20850, USA

Tel: 301-838-0200

Fax: 301-838-0208

Email: ekirknes@tigr.org

Class: shotgun.

Location/Qualifiers

1. .700

/organism="Canis familiaris"

/mol\_type="genomic DNA"

/strain="Standard Poodle"

/db\_xref="taxon:9615"

## FEATURES

source

/clone\_lib="Dog Library"  
/note="Site 1: BstXI; Libraries were prepared from  
peripheral blood"

## ORIGIN

Query Match 8.3%; Score 39.6; DB 29; Length 700;  
Best Local Similarity 58.5%; Pred. No. 82;  
Matches 69; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 351 CTCACATTTAGCTTTAATAATATGTTTCATTAATTTCTGTTTGTCTCTATATA 410

Db 644 CTTTAAATTTGGGTTTATCTAATGTTCTTATGATTAGATGTTCTGCAATTTCTTATA 585

QY 411 AAGAGCTTGTATTTTCATTTGAAGCAGAGCGGCAACACACACAGAACCTCCCTGCT 468

Db 584 AAAATGTGTGTTTCTTATGATGCAATTTAAATCAGACATAGAGGCACCTGGGT 527

## RESULT 14

AL544517/c

LOCUS

DEFINITION AL544517 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS0D1021Y122 5-PRIME, mRNA sequence.

ACCESSION AL544517

VERSION AL544517.2 GI:31266360

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1201)

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On Feb 15, 2001 this sequence version replaced gi:12876997.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 3154.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0D1021BE11Q1&cluster=3154.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0D1021BE11Q1.

Location/Qualifiers

1. 1201

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D1021Y122"

/tissue type="PLACENTA"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dT)

primer. Five prime end enriched, double-strand cDNA was

digested with NotI and cloned into the NotI and EcoRV

sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 8.3%; Score 39.6; DB 9; Length 1201;  
Best Local Similarity 31.9%; Pred. No. 76;  
Matches 102; Conservative 58; Mismatches 160; Indels 0; Gaps 0;

QY 122 AAAGACAGCAAGACATGGTGTCTCGATGCACAGAACACATCTTTGACGAGATGTG 181

Db 1117 MWMWMMWMMWMAAAANDWMMWYCCCGGGSSNAAMCCCGGGKKKKKBYCAAAWWTY 1058

QY 182 AAGCAGCCAGATGTGTCACAGAGCGACTCAGAAAGAGGCATCTTACCCACACAGAAA 241

Db 1057 CCCCCVAAAAAARRGGKCCYCAAAAAAAMCVAATAAATTTTCCCHCMWRRDITWAA 998



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 08:52:25 ; Search time 2305 Seconds  
(without alignments)  
8950.666 Million cell updates/sec

Title: US-10-020-540A-1

Perfect score: 476

Sequence: 1 aactatgctgatgacaagat.....aacctccctgcttacaacc 476

Scoring table:

OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.ste.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.man.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	476	100.0	476	6	AX474089	AX474089 Sequence
2	476	100.0	518	6	AX474090	AX474090 Sequence
3	476	100.0	7876	14	SVBVCOMGN	X97304 Strawberry
4	435	91.4	950	14	AF331666	AF331666 Strawberry
5	369	77.5	651	6	AX474091	AX474091 Sequence
6	369	77.5	701	6	AX474092	AX474092 Sequence
c 7	21	4.4	41339	9	AC005490	AC005490 Homo sapi
c 8	21	4.4	192672	2	AC121285	AC121285 Mus muscu
c 9	21	4.4	221190	2	BX571704	AX474090 Sequence
c 10	20	4.2	1046	8	AK061232	AX474090 Sequence
c 11	20	4.2	1490	8	SCYKL041W	AX474090 Sequence
c 12	20	4.2	1557	6	AX008583	AX474090 Sequence
c 13	20	4.2	1804	8	SCYKL040C	AX474090 Sequence
c 14	20	4.2	2204	8	AF307842	AX474090 Sequence
c 15	20	4.2	2709	10	BC051144	BC051144 Mus muscu
c 16	20	4.2	2924	10	BC021479	BC021479 Mus muscu
c 17	20	4.2	6576	8	YSCSEC7	YSCSEC7 Yeast (S.ce
c 18	20	4.2	9539	8	AF307843	AF307843 Chlamydom
c 19	20	4.2	13631	8	SC9489	SC9489 S.cerevisia
c 20	20	4.2	17564	8	SC9489	SC9489 S.cerevisia
c 21	20	4.2	17594	9	AF003093	AF003093 Homo sapi
c 22	20	4.2	83007	8	AC123575	AC123575 Medicago
c 23	20	4.2	86314	9	AC109592	AC109592 Homo sapi
c 24	20	4.2	106273	9	HSJ1018A4	HSJ1018A4 Human DNA
c 25	20	4.2	110000	2	AC111399-2	Continuation (3 of
c 26	20	4.2	141529	2	AC025330	AC025330 Homo sapi
c 27	20	4.2	151494	2	AC132889	AC132889 Rattus no
c 28	20	4.2	156945	10	AC105944	AC105944 Mus muscu
c 29	20	4.2	159857	14	AF170722	AF170722 Rabbitt fi
c 30	20	4.2	160847	8	AC120307	AC120307 Oryza sat
c 31	20	4.2	163120	9	AC011454	AC011454 Homo sapi
c 32	20	4.2	171493	2	AC067933	AC067933 Homo sapi
c 33	20	4.2	173967	9	AC073343	AC073343 Homo sapi
c 34	20	4.2	174526	2	AC101807	AC101807 Mus muscu
c 35	20	4.2	176026	9	AC092661	AC092661 Homo sapi
c 36	20	4.2	184506	10	AC112701	AC112701 Mus muscu
c 37	20	4.2	188412	2	AC119098	AC119098 Rattus no
c 38	20	4.2	192843	10	AL607124	AL607124 Mouse DNA
c 39	20	4.2	197352	2	AC103365	AC103365 Mus muscu
c 40	20	4.2	205621	10	AC079082	AC079082 Mus muscu
c 41	20	4.2	207922	10	AL450399	AL450399 Mouse DNA
c 42	20	4.2	213875	10	AC096975	AC096975 Rattus no
c 43	20	4.2	214831	2	AC125984	AC125984 Rattus no
c 44	20	4.2	220446	2	AC129443	AC129443 Rattus no
c 45	20	4.2	223979	2	AC096151	AC096151 Rattus no

# ALIGNMENTS

RESULT 1	AX474089	476 bp	DNA	linear	PAT 09-AUG-2002
LOCUS	Sequence 1 from Patent WO0240691.				
DEFINITION	AX474089				
ACCESSION	AX474089.1	GI:22208238			
VERSION	AX474089.1				
KEYWORDS	Strawberry vein banding virus (SVBV)				
SOURCE	Strawberry vein banding virus				
ORGANISM	Viruses; Retroviridae; Caulimoviridae; Caulimovirus.				
REFERENCE	1				
AUTHORS	Wu, G. and McBride, K.				
TITLE	Strawberry vein banding virus (svbv) promoter				
JOURNAL	Patent: WO 0240691-A 1 23-MAY-2002;				
	Maxygen, Inc. (US)				

FEATURES		Location/Qualifiers
source		1..476
		/organism="Strawberry vein banding virus"
		/mol_type="unassigned DNA"
		/strain="E3"
		/db_xref="taxon:47903"
ORIGIN		
Query Match		100.0%; Score 476; DB 6; Length 476;
Best Local Similarity		100.0%; Pred. No. 1.4e-261;
Matches 476; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
QY	1	AACATGCTGATGACAAAGATAATTTCTAATAGCAATTAATTCAGAAATTAATCAAGAGAAA 60
DB	1	AACATGCTGATGACAAAGATAATTTCTAATAGCAATTAATTCAGAAATTAATCAAGAGAAA 60
QY	61	GAATTAATAACTCTTTTCAGAAATATGAAGCCCGCTTTTACAAGTGGCCAGCTAGCTATCACT 120
DB	61	GAATTAATAACTCTTTTCAGAAATATGAAGCCCGCTTTTACAAGTGGCCAGCTAGCTATCACT 120
QY	121	GAAAGACAGCAGCAATGCTGCTCGATGACCAAGAACCAACATCTTTGACGAGATGT 180
DB	121	GAAAGACAGCAGCAATGCTGCTCGATGACCAAGAACCAACATCTTTGACGAGATGT 180
QY	181	GAAAGACAGCAGCAATGCTGCTCGATGACCAAGAACCAACATCTTTGACGAGATGT 240
DB	181	GAAAGACAGCAGCAATGCTGCTCGATGACCAAGAACCAACATCTTTGACGAGATGT 240
QY	241	GAATTAATAACTCTTTTCAGAAATATGAAGCCCGCTTTTACAAGTGGCCAGCTAGCTATCACT 300
DB	241	GAATTAATAACTCTTTTCAGAAATATGAAGCCCGCTTTTACAAGTGGCCAGCTAGCTATCACT 300
QY	301	AGACTGACCCAGGCGGACACTTAAGAGAAATATGCAAGTGGCTTCCACTTTA 360
DB	301	AGACTGACCCAGGCGGACACTTAAGAGAAATATGCAAGTGGCTTCCACTTTA 360
QY	361	GCTTTAATAATTAATTTTCAATTAATTTCTGCTTTTCTCTATATAAAGAGCTTGT 420
DB	361	GCTTTAATAATTAATTTTCAATTAATTTCTGCTTTTCTCTATATAAAGAGCTTGT 420
QY	421	ATTTCATTGAGCGAGCGGACACACACACAGACCTCCCTGCTTACAAACC 476
DB	421	ATTTCATTGAGCGAGCGGACACACACACAGACCTCCCTGCTTACAAACC 476
RESULT 2		
AX474090		
LOCUS		518 bp DNA linear PAT 09-AUG-2002
DEFINITION		Sequence 2 from Patent WO0240691.
ACCESSION		AX474090
VERSION		AX474090.1 GI:22208239
KEYWORDS		Strawberry vein banding virus (SVBV)
SOURCE		Strawberry vein banding virus
ORGANISM		Viruses; Retroviridae; Caulimoviridae; Caulimovirus.
REFERENCE		1 Wu, G. and McBride, K.
AUTHORS		Strawberry vein banding virus (svbv) promoter
TITLE		Patent: WO 0240691-A 2 23-MAY-2002;
JOURNAL		Maxygen, Inc. (US)
FEATURES		
source		1..518
		/organism="Strawberry vein banding virus"
		/mol_type="unassigned DNA"
		/strain="E3"
		/db_xref="taxon:47903"
ORIGIN		
Query Match		100.0%; Score 476; DB 6; Length 518;
Best Local Similarity		100.0%; Pred. No. 1.4e-261;
Matches 476; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
QY	1	AACATGCTGATGACAAAGATAATTTCTAATAGCAATTAATTCAGAAATTAATCAAGAGAAA 60
FEATURES		
source		1..7876
		/organism="Strawberry vein banding virus"
		/mol_type="genomic DNA"
		/specific_host="Fragaria sp."
		/db_xref="taxon:47903"
		/clone="pSVBV-E3"
		/gene="ORF I"
		67..1053
		/gene="ORF I"
		67..1053
		/codon_start=1
		/protein_id="CAA65972.1"
		/db_xref="GI:1360609"
		/db_xref="GOA:Q88438"
		/db_xref="SPREMBL:Q88438"
		/translation="MSEEEIRMDQGGHDBYIFEEETVAHDVAIDSTLLKEIEKKD"
RESULT 3		
SVBVCOMGN		
LOCUS		7876 bp DNA linear VRL 20-JUL-2001
DEFINITION		Strawberry vein banding virus complete genome.
ACCESSION		X97304
VERSION		X97304.1 GI:1360608
KEYWORDS		complete genome; ORFI; ORFII; ORFIII; orfIV; ORFV; ORFVI.
SOURCE		Strawberry vein banding virus (SVBV)
ORGANISM		Strawberry vein banding virus
REFERENCE		1 Petrzik, K., Benes, V., Mraz, I., Honetslegrova-Franova, J., Ansoerge, W. and Spak, J.
AUTHORS		Strawberry vein banding virus--definitive member of the genus Caulimovirus
TITLE		Strawberry vein banding virus--definitive member of the genus Caulimovirus
JOURNAL		Virus Genes 16 (3), 303-305 (1998)
MEDLINE		98318753
PUBMED		9654684
REFERENCE		2 (bases 1 to 7876)
AUTHORS		Petrzik, K.
TITLE		Direct Submission
JOURNAL		Submitted (03-APR-1996) K. Petrzik, Institute of Plant Molecular Biology, Dept. Plant Virology, Branisovska 31, Ceske Budejovice, CZ-370 05, Czech Republic
FEATURES		
Location/Qualifiers		
source		1..7876
		/organism="Strawberry vein banding virus"
		/mol_type="genomic DNA"
		/specific_host="Fragaria sp."
		/db_xref="taxon:47903"
		/clone="pSVBV-E3"
		/gene="ORF I"
		67..1053
		/gene="ORF I"
		67..1053
		/codon_start=1
		/protein_id="CAA65972.1"
		/db_xref="GI:1360609"
		/db_xref="GOA:Q88438"
		/db_xref="SPREMBL:Q88438"
		/translation="MSEEEIRMDQGGHDBYIFEEETVAHDVAIDSTLLKEIEKKD"

```
gene
CDS
LELTSEVFKTSLMKKFLKARKNICISCVSREYPIEIAQANGLETPFFNREBIES
KXVLKPERKIDFTHIGSVKIMIKSTRTIGDPSISVALLDRMKNAKDAVFGVK
GNLSYKLFITNPKISVSLRDPDTINKTLTLAHFEKEKELMEHGNHPVTISYKIVTL
SNSHSLERPRKPCIDDLFTSVGKISQAPLOEITPIENYWRMSLGNSTRLLGERP
RMVIEDDEETPLDRHETARALARSORMLGRPLRDLRTTKRKINILAHQL"
1058. 1543
/gene="ORF II"
1058. 1543
/gene="ORF II"
/codon_start=1
/protein_id="CAA65973.1"
/db_xref="GI:1360610"
/translation="MSFRREPHIYKKQVLTNTQNVGERPEYLYVCGRGIEGLKH
LNNIVVCGNIHMNYIKTIGIKTSIYRNPEDNSPGLSLILGKPOESSFONPLE
KKIDNLSKIDRLGNSNTTGTETEKLADLSNKISELERKISSLTDDINMLLKLDK
C"
1543. 1896
/gene="ORF III"
1543. 1896
/gene="ORF III"
/codon_start=1
/protein_id="CAA65974.1"
/db_xref="GI:1360611"
/db_xref="SPTREMBL:Q88440"
/translation="MGSVTKTERIYDEMVKMEERSRLERQEQDHFSEKIDQASAIL
LAEIRKIQSKLETCDCKEILDAKQDRTKKNGSPSNLDAALDLAHKONSPRNPRI
GTENWHPDQLKLTW"
1890. 3314
/gene="ORF IV"
1890. 3314
/gene="ORF IV"
/codon_start=1
/protein_id="CAA65969.1"
/db_xref="GI:1360612"
/db_xref="GOA:Q88441"
/db_xref="SPTREMBL:Q88441"
/translation="MYSVRERLEOLEDEHPMDIIOVLSLDELDCQEKLOVL
VAKEMIDYSSDEDETRNVKVIKEDEETVRPNKRKRGNSSSQNYRYRIPQEI
PNRKTGLNTSLDLDCASNNRLOIBEWNDNEMLIITKEKALTNDLTLTAKKTV
GNAQLLESHTFAFQAOSTGEEFLTITNSFYTFVGTNYLTQOTREKXAVGAR
NRVLQICNLHSPFCDDYETNKLPIEEMPKYIEBYIRKIPVGMVELEYSKO
DSTTKGSLVGNHLIKAYEKKCKSLIKKEIRNMCCPKFSPETQYCKGKPISHKA
KKQKQYKVKYKRLKPKBWTNSRRKYSRKLFRFRKQDKSETSOQSPBEKKKFCPO
KVTCTCWINCHGPAKDCRNKNSAHNKKIIELOSLQLEVPVDELNELKIEKFHELK
EVSSESESEISDESSESDLE"
3406. 5532
/gene="ORF V"
3406. 5532
/gene="ORF V"
/codon_start=1
/protein_id="CAA65970.1"
/db_xref="GI:1360613"
/db_xref="GOA:Q88442"
/db_xref="SPTREMBL:Q88442"
/translation="MENNQLLELHLEEGWMLTDEDFKLOONDEBQLLESVDSE
ERPFLTNTPNSTYIRGNFYFKYKSLDLVDVTGASCTANKHVIPEEFWNAKN
PIRARIANDNFKVAKELMQVQIADETIIPLYCATTKGDTIGNNFCEVFPV
QYKDMITPHDGRAVSKYTKAYFHLPGFLESKKVSGSTSTPNENTPTVINGQDN
ISKIFGEEIGESEQLFSTISAYTEVEKLDSICSEHPLDSRINKGRFQAALLDTN
KITCKPMQSPQDREFFKTQIEELKLGIIRPSKSPHSPAFMVRNHAIRKRGKRN
VINYKLNHDTKGDYLLPNKQLOLRIGKTFYSFDCSKSFWQVRLAPETIQLTAF
SCPOGHYELMWLPFGFLKPAIFORHMDLSNNYPOFCAYVYVDIIVFSKTEEHLG
HVKTIVNRKALGVLSKKAOLCKTINFLGLVIBRGNLKVQSHIGLHLVAPPDOLS
DRNALORPLGLMVISAYFPKIANLSPLQVOKLKEITWSWTEKDTETVRKIKSLVKT
LPDNLSPDKPFIICDASDDHWGAILKALPEGKEVICRYASGTFKPAEKNYHNE
KEILUSILKAFKAYILPYKFLVTRNTAAVFRNTIAGDQKSLRVRWQNALREY
SFDIEHVSQGNVLADIMTRELAKT"
5544. 7088
/gene="ORF VI"
5544. 7088
/gene="ORF VI"
/codon_start=1
```

```
ORIGIN
Query Match 100.0%; Score 476; DB 14; Length 7876;
Best Local Similarity 100.0%; Pred. No. 1.5e-261; Indels 0; Gaps 0;
Matches 476; Conservative 0; Mismatches 0;

QY 1 AACTATGCTGATGACAGATAATTCTTAATAAGCAATTATTAGAAATTAATCAAGAGAAA 60
DB 6816 AACTATGCTGATGACAGATAATTCTTAATAAGCAATTATTAGAAATTAATCAAGAGAAA 6875
QY 61 GAATTAATAACTCTTTCAGRATATGAAGCCGCTTCAAGTGGCCAGCTAGCTATCACT 120
DB 6876 GAATTAATAACTCTTTCAGRATATGAAGCCGCTTCAAGTGGCCAGCTAGCTATCACT 6935
QY 121 GAAAACAGACAGACAATGCTGCTCGATGACACAGAACACACATCTTTGACGAGATGT 180
DB 6936 GAAAACAGACAGACAATGCTGCTCGATGACACAGAACACACATCTTTGACGAGATGT 6995
QY 181 GAAGCAGCAGAGTGCTCCACAGAGCGCATCAGAAAAGGATCTTCTACGACACAGAA 240
DB 6996 GAAGCAGCAGAGTGCTCCACAGAGCGCATCAGAAAAGGATCTTCTACGACACAGAA 7055
QY 241 AAAGCAACACAGCTCATCATCAACATGATAGCTGCTTATCGGTCGGCTGAAGATA 300
DB 7056 AAAGCAACACAGCTCATCATCAACATGATAGCTGCTTATCGGTCGGCTGAAGATA 7115
QY 301 AGACTGACCCAGGCGCAGCATTAAGAGAAATAATCAAGTGGCTAGCTCCACTTTA 360
DB 7116 AGACTGACCCAGGCGCAGCATTAAGAGAAATAATCAAGTGGCTAGCTCCACTTTA 7175
QY 361 GCTTTTAATAATATGCTTTTCAATATTTCTGCTTTTGTCTCTATATAAAGAGCTTGT 420
DB 7176 GCTTTTAATAATATGCTTTTCAATATTTCTGCTTTTGTCTCTATATAAAGAGCTTGT 7235
QY 421 ATTTTCATTGTAAGGCGAGGCGGAACACACACAGAACCTCCCTGCTTACAAACC 476
DB 7236 ATTTTCATTGTAAGGCGAGGCGGAACACACACAGAACCTCCCTGCTTACAAACC 7291

RESULT 4
AF331666
LOCUS
DEFINITION
Strawberry vein banding virus ORF VI (ORF VI) gene, partial cds and
viral promoter sequence.
ACCESSION
AF331666
VERSION
AF331666.1 GI:13345788
KEYWORDS
Strawberry vein banding virus
Strawberry vein banding virus
Viruses; Retroid viruses; Caulimoviridae; Caulimovirus.
ORGANISM
1 (bases 1 to 950)
Wang, Y., Gaba, V., Wolf, D., Xia, X.D., Zelcer, A. and Gal-On, A.
Identification of a novel plant virus promoter using a potyvirus
infectious clone
JOURNAL
Virus Genes 20 (1), 11-17 (2000)
MEDLINE
2027369
PUBMED
10766302
REFERENCE
2 (bases 1 to 950)
Wang, Y., Gaba, V., Wolf, D., Xia, X.D., Zelcer, A. and Gal-On, A.
```



TITLE Direct Submission  
JOURNAL Submitted (21-DEC-2000) Plant Virology Dept., The Volcani  
Center-ARO, Bet Dagan 50-250, Israel  
FEATURES Location/Qualifiers  
source 1..950  
/organism="Strawberry vein banding virus"  
/proviral  
/mol\_type="genomic DNA"  
/db\_xref="taxon:47903"  
c1..788  
/gene="ORF VI"  
c1..788  
/gene="ORF VI"  
/codon\_start=3  
/product="ORF VI"  
/protein\_id="AAK19514.1"  
/db\_xref="GI:13345789"  
/translation="PLVTKDHFHNYLVAQTHDNCBEKFTFVDSKTLISYGFIEGAD  
PEYVREAFHCGLSLIPYFENLKEISLPKOLYRAIKNTRVCSQSRSPIYLRTST  
FLEWENDILIPYHWEIYGLSNGKIPMDKSYKKTATVEYLPETALRPKRIKAVYNO  
TKIFADSKKLYNADDKILISNYSELTKEKELITLGEYEARFTSGQLAITEKTAQW  
CLDAPPEPFLCSECAARVVVHTHSEKASSTDEKDNHSSSNM"  
promoter 523..950  
/note="sufficient to act as a viral promoter"  
ORIGIN  
Query Match 91.4%; Score 435; DB 14; Length 950;  
Best Local Similarity 100.0%; Pred. No. 5.1e-238;  
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AACTATGCTGATGACAAAGATAATTTCTAATAGCAATTTATCAGAAATTAATCAAGAGAAA 60  
DB 516 AACTATGCTGATGACAAAGATAATTTCTAATAGCAATTTATCAGAAATTAATCAAGAGAAA 575  
QY 61 GAATTAATAACTCTTTTCAAGATATGACGCGCTTTTACAAGTGGCCAGCTAGCTATCACT 120  
DB 576 GAATTAATAACTCTTTTCAAGATATGACGCGCTTTTACAAGTGGCCAGCTAGCTATCACT 635  
QY 121 GAAAGACAGCAGAGCAATGGTGTCTCGATGACACAGAACCAATCTTTTGACGAGATGT 180  
DB 636 GAAAGACAGCAGAGCAATGGTGTCTCGATGACACAGAACCAATCTTTTGACGAGATGT 695  
QY 181 GAAGCAGCAGAGTGGTCCACAGAGCGCACTCAGAAAGGATCTTCTACCGACACAGAA 240  
DB 696 GAAGCAGCAGAGTGGTCCACAGAGCGCACTCAGAAAGGATCTTCTACCGACACAGAA 755  
QY 241 AAAGACACACAGCTCATCTCCAAACATGTAGACTGCTGTTATGCTGCGCTGAAGATA 300  
DB 756 AAAGACACACAGCTCATCTCCAAACATGTAGACTGCTGTTATGCTGCGCTGAAGATA 815  
QY 301 AGACTGACCCAGCCAGCAGCACTAAAGAGAAATAATGCAAGTGGTCTAGCTCCACTTTA 360  
DB 816 AGACTGACCCAGCCAGCAGCACTAAAGAGAAATAATGCAAGTGGTCTAGCTCCACTTTA 875  
QY 361 GCTTTAATAATATGTTTCATTATTTATTTCTGCTTTTCTCTATATAAGAGCTTGT 420  
DB 876 GCTTTAATAATATGTTTCATTATTTATTTCTGCTTTTCTCTATATAAGAGCTTGT 935  
QY 421 ATTTTCATTGGAAGG 435  
DB 936 ATTTTCATTGGAAGG 950  
RESULT 5  
LOCUS AX474091 651 bp DNA linear PAT 09-AUG-2002  
DEFINITION Sequence 3 from Patent WO0240691.  
ACCESSION AX474091  
VERSION AX474091.1 GI:22208240  
KEYWORDS  
SOURCE Strawberry vein banding virus (SVBV)  
ORGANISM Strawberry vein banding virus (SVBV)  
Viruses; Retroviridae; Caulimoviridae; Caulimovirus.

REFERENCE 1  
AUTHORS Wu, G. and McBride, K.  
TITLE Strawberry vein banding virus (svbv) promoter  
JOURNAL Patent: WO 0240691-A 3 23-MAY-2002;  
Maxygen, Inc. (US)  
FEATURES Location/Qualifiers  
source 1..651  
/organism="Strawberry vein banding virus"  
/mol\_type="unassigned DNA"  
/strain="E3"  
/db\_xref="taxon:47903"  
ORIGIN  
Query Match 77.5%; Score 369; DB 6; Length 651;  
Best Local Similarity 100.0%; Pred. No. 4.2e-200;  
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 108 GCTAGCTATCACTGAAAAGACAGCAAGCAATGGTGTCTCGATGCAACCAAGCAATCT 167  
DB 278 GCTAGCTATCACTGAAAAGACAGCAAGCAATGGTGTCTCGATGCAACCAAGCAATCT 337  
QY 168 TTCCACAGATGTGAAGCAGCCAGAGTGTCCACAGAGCGCACTCAGAAAAGGCATCTTC 227  
DB 338 TTCCACAGATGTGAAGCAGCCAGAGTGTCCACAGAGCGCACTCAGAAAAGGCATCTTC 397  
QY 228 TACCGACACAGAAAAGACACACAGCTCATCTCCAAACATGTAGACTGTCTTATGCG 287  
DB 398 TACCGACACAGAAAAGACACACAGCTCATCTCCAAACATGTAGACTGTCTTATGCG 457  
QY 288 TCGGCTGAAGATAAGACTGACCCGAGCCAGCACTAAAGAGAAATAATGCAAGTGGTCC 347  
DB 458 TCGGCTGAAGATAAGACTGACCCGAGCCAGCACTAAAGAGAAATAATGCAAGTGGTCC 517  
QY 348 TAGCTCCACTTTAGCTTTAATTAATTTATGTTTCAATTTATTTCTGCTTTTCTCTAT 407  
DB 518 TAGCTCCACTTTAGCTTTAATTAATTTATGTTTCAATTTATTTCTGCTTTTCTCTAT 577  
QY 408 ATAAAGAGCTGTATTTTCAATTTGAAGCAGAGGGGAAACACACACAGAACCTCCCTGC 467  
DB 578 ATAAAGAGCTGTATTTTCAATTTGAAGCAGAGGGGAAACACACACAGAACCTCCCTGC 637  
QY 468 TTACAAACC 476  
DB 638 TTACAAACC 646  
RESULT 6  
LOCUS AX474092 701 bp DNA linear PAT 09-AUG-2002  
DEFINITION Sequence 4 from Patent WO0240691.  
ACCESSION AX474092  
VERSION AX474092.1 GI:22208241  
KEYWORDS  
SOURCE Strawberry vein banding virus (SVBV)  
ORGANISM Strawberry vein banding virus (SVBV)  
Viruses; Retroviridae; Caulimoviridae; Caulimovirus.  
REFERENCE 1  
AUTHORS Wu, G. and McBride, K.  
TITLE Strawberry vein banding virus (svbv) promoter  
JOURNAL Patent: WO 0240691-A 4 23-MAY-2002;  
Maxygen, Inc. (US)  
FEATURES Location/Qualifiers  
source 1..701  
/organism="Strawberry vein banding virus"  
/mol\_type="unassigned DNA"  
/strain="E3"  
/db\_xref="taxon:47903"  
ORIGIN  
Query Match 77.5%; Score 369; DB 6; Length 701;  
Best Local Similarity 100.0%; Pred. No. 4.2e-200;  
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



108 GCTAGTATCACTGAAACACAGCAGCAATGGTGTCTCGATGCACAGAACCATCT 167  
109 |||||  
110 |||||  
111 |||||  
112 |||||  
113 |||||  
114 |||||  
115 |||||  
116 |||||  
117 |||||  
118 |||||  
119 |||||  
120 |||||  
121 |||||  
122 |||||  
123 |||||  
124 |||||  
125 |||||  
126 |||||  
127 |||||  
128 |||||  
129 |||||  
130 |||||  
131 |||||  
132 |||||  
133 |||||  
134 |||||  
135 |||||  
136 |||||  
137 |||||  
138 |||||  
139 |||||  
140 |||||  
141 |||||  
142 |||||  
143 |||||  
144 |||||  
145 |||||  
146 |||||  
147 |||||  
148 |||||  
149 |||||  
150 |||||  
151 |||||  
152 |||||  
153 |||||  
154 |||||  
155 |||||  
156 |||||  
157 |||||  
158 |||||  
159 |||||  
160 |||||  
161 |||||  
162 |||||  
163 |||||  
164 |||||  
165 |||||  
166 |||||  
167 |||||  
168 |||||  
169 |||||  
170 |||||  
171 |||||  
172 |||||  
173 |||||  
174 |||||  
175 |||||  
176 |||||  
177 |||||  
178 |||||  
179 |||||  
180 |||||  
181 |||||  
182 |||||  
183 |||||  
184 |||||  
185 |||||  
186 |||||  
187 |||||  
188 |||||  
189 |||||  
190 |||||  
191 |||||  
192 |||||  
193 |||||  
194 |||||  
195 |||||  
196 |||||  
197 |||||  
198 |||||  
199 |||||  
200 |||||  
201 |||||  
202 |||||  
203 |||||  
204 |||||  
205 |||||  
206 |||||  
207 |||||  
208 |||||  
209 |||||  
210 |||||  
211 |||||  
212 |||||  
213 |||||  
214 |||||  
215 |||||  
216 |||||  
217 |||||  
218 |||||  
219 |||||  
220 |||||  
221 |||||  
222 |||||  
223 |||||  
224 |||||  
225 |||||  
226 |||||  
227 |||||  
228 |||||  
229 |||||  
230 |||||  
231 |||||  
232 |||||  
233 |||||  
234 |||||  
235 |||||  
236 |||||  
237 |||||  
238 |||||  
239 |||||  
240 |||||  
241 |||||  
242 |||||  
243 |||||  
244 |||||  
245 |||||  
246 |||||  
247 |||||  
248 |||||  
249 |||||  
250 |||||  
251 |||||  
252 |||||  
253 |||||  
254 |||||  
255 |||||  
256 |||||  
257 |||||  
258 |||||  
259 |||||  
260 |||||  
261 |||||  
262 |||||  
263 |||||  
264 |||||  
265 |||||  
266 |||||  
267 |||||  
268 |||||  
269 |||||  
270 |||||  
271 |||||  
272 |||||  
273 |||||  
274 |||||  
275 |||||  
276 |||||  
277 |||||  
278 |||||  
279 |||||  
280 |||||  
281 |||||  
282 |||||  
283 |||||  
284 |||||  
285 |||||  
286 |||||  
287 |||||  
288 |||||  
289 |||||  
290 |||||  
291 |||||  
292 |||||  
293 |||||  
294 |||||  
295 |||||  
296 |||||  
297 |||||  
298 |||||  
299 |||||  
300 |||||  
301 |||||  
302 |||||  
303 |||||  
304 |||||  
305 |||||  
306 |||||  
307 |||||  
308 |||||  
309 |||||  
310 |||||  
311 |||||  
312 |||||  
313 |||||  
314 |||||  
315 |||||  
316 |||||  
317 |||||  
318 |||||  
319 |||||  
320 |||||  
321 |||||  
322 |||||  
323 |||||  
324 |||||  
325 |||||  
326 |||||  
327 |||||  
328 |||||  
329 |||||  
330 |||||  
331 |||||  
332 |||||  
333 |||||  
334 |||||  
335 |||||  
336 |||||  
337 |||||  
338 |||||  
339 |||||  
340 |||||  
341 |||||  
342 |||||  
343 |||||  
344 |||||  
345 |||||  
346 |||||  
347 |||||  
348 |||||  
349 |||||  
350 |||||  
351 |||||  
352 |||||  
353 |||||  
354 |||||  
355 |||||  
356 |||||  
357 |||||  
358 |||||  
359 |||||  
360 |||||  
361 |||||  
362 |||||  
363 |||||  
364 |||||  
365 |||||  
366 |||||  
367 |||||  
368 |||||  
369 |||||  
370 |||||  
371 |||||  
372 |||||  
373 |||||  
374 |||||  
375 |||||  
376 |||||  
377 |||||  
378 |||||  
379 |||||  
380 |||||  
381 |||||  
382 |||||  
383 |||||  
384 |||||  
385 |||||  
386 |||||  
387 |||||  
388 |||||  
389 |||||  
390 |||||  
391 |||||  
392 |||||  
393 |||||  
394 |||||  
395 |||||  
396 |||||  
397 |||||  
398 |||||  
399 |||||  
400 |||||  
401 |||||  
402 |||||  
403 |||||  
404 |||||  
405 |||||  
406 |||||  
407 |||||  
408 |||||  
409 |||||  
410 |||||  
411 |||||  
412 |||||  
413 |||||  
414 |||||  
415 |||||  
416 |||||  
417 |||||  
418 |||||  
419 |||||  
420 |||||  
421 |||||  
422 |||||  
423 |||||  
424 |||||  
425 |||||  
426 |||||  
427 |||||  
428 |||||  
429 |||||  
430 |||||  
431 |||||  
432 |||||  
433 |||||  
434 |||||  
435 |||||  
436 |||||  
437 |||||  
438 |||||  
439 |||||  
440 |||||  
441 |||||  
442 |||||  
443 |||||  
444 |||||  
445 |||||  
446 |||||  
447 |||||  
448 |||||  
449 |||||  
450 |||||  
451 |||||  
452 |||||  
453 |||||  
454 |||||  
455 |||||  
456 |||||  
457 |||||  
458 |||||  
459 |||||  
460 |||||  
461 |||||  
462 |||||  
463 |||||  
464 |||||  
465 |||||  
466 |||||  
467 |||||  
468 |||||  
469 |||||  
470 |||||  
471 |||||  
472 |||||  
473 |||||  
474 |||||  
475 |||||  
476 |||||  
477 |||||  
478 |||||  
479 |||||  
480 |||||  
481 |||||  
482 |||||  
483 |||||  
484 |||||  
485 |||||  
486 |||||  
487 |||||  
488 |||||  
489 |||||  
490 |||||  
491 |||||  
492 |||||  
493 |||||  
494 |||||  
495 |||||  
496 |||||  
497 |||||  
498 |||||  
499 |||||  
500 |||||  
501 |||||  
502 |||||  
503 |||||  
504 |||||  
505 |||||  
506 |||||  
507 |||||  
508 |||||  
509 |||||  
510 |||||  
511 |||||  
512 |||||  
513 |||||  
514 |||||  
515 |||||  
516 |||||  
517 |||||  
518 |||||  
519 |||||  
520 |||||  
521 |||||  
522 |||||  
523 |||||  
524 |||||  
525 |||||  
526 |||||  
527 |||||  
528 |||||  
529 |||||  
530 |||||  
531 |||||  
532 |||||  
533 |||||  
534 |||||  
535 |||||  
536 |||||  
537 |||||  
538 |||||  
539 |||||  
540 |||||  
541 |||||  
542 |||||  
543 |||||  
544 |||||  
545 |||||  
546 |||||  
547 |||||  
548 |||||  
549 |||||  
550 |||||  
551 |||||  
552 |||||  
553 |||||  
554 |||||  
555 |||||  
556 |||||  
557 |||||  
558 |||||  
559 |||||  
560 |||||  
561 |||||  
562 |||||  
563 |||||  
564 |||||  
565 |||||  
566 |||||  
567 |||||  
568 |||||  
569 |||||  
570 |||||  
571 |||||  
572 |||||  
573 |||||  
574 |||||  
575 |||||  
576 |||||  
577 |||||  
578 |||||  
579 |||||  
580 |||||  
581 |||||  
582 |||||  
583 |||||  
584 |||||  
585 |||||  
586 |||||  
587 |||||  
588 |||||  
589 |||||  
590 |||||  
591 |||||  
592 |||||  
593 |||||  
594 |||||  
595 |||||  
596 |||||  
597 |||||  
598 |||||  
599 |||||  
600 |||||  
601 |||||  
602 |||||  
603 |||||  
604 |||||  
605 |||||  
606 |||||  
607 |||||  
608 |||||  
609 |||||  
610 |||||  
611 |||||  
612 |||||  
613 |||||  
614 |||||  
615 |||||  
616 |||||  
617 |||||  
618 |||||  
619 |||||  
620 |||||  
621 |||||  
622 |||||  
623 |||||  
624 |||||  
625 |||||  
626 |||||  
627 |||||  
628 |||||  
629 |||||  
630 |||||  
631 |||||  
632 |||||  
633 |||||  
634 |||||  
635 |||||  
636 |||||  
637 |||||  
638 |||||  
639 |||||  
640 |||||  
641 |||||  
642 |||||  
643 |||||  
644 |||||  
645 |||||  
646 |||||  
647 |||||  
648 |||||  
649 |||||  
650 |||||  
651 |||||  
652 |||||  
653 |||||  
654 |||||  
655 |||||  
656 |||||  
657 |||||  
658 |||||  
659 |||||  
660 |||||  
661 |||||  
662 |||||  
663 |||||  
664 |||||  
665 |||||  
666 |||||  
667 |||||  
668 |||||  
669 |||||  
670 |||||  
671 |||||  
672 |||||  
673 |||||  
674 |||||  
675 |||||  
676 |||||  
677 |||||  
678 |||||  
679 |||||  
680 |||||  
681 |||||  
682 |||||  
683 |||||  
684 |||||  
685 |||||  
686 |||||  
687 |||||  
688 |||||  
689 |||||  
690 |||||  
691 |||||  
692 |||||  
693 |||||  
694 |||||  
695 |||||  
696 |||||  
697 |||||  
698 |||||  
699 |||||  
700 |||||  
701 |||||  
702 |||||  
703 |||||  
704 |||||  
705 |||||  
706 |||||  
707 |||||  
708 |||||  
709 |||||  
710 |||||  
711 |||||  
712 |||||  
713 |||||  
714 |||||  
715 |||||  
716 |||||  
717 |||||  
718 |||||  
719 |||||  
720 |||||  
721 |||||  
722 |||||  
723 |||||  
724 |||||  
725 |||||  
726 |||||  
727 |||||  
728 |||||  
729 |||||  
730 |||||  
731 |||||  
732 |||||  
733 |||||  
734 |||||  
735 |||||  
736 |||||  
737 |||||  
738 |||||  
739 |||||  
740 |||||  
741 |||||  
742 |||||  
743 |||||  
744 |||||  
745 |||||  
746 |||||  
747 |||||  
748 |||||  
749 |||||  
750 |||||  
751 |||||  
752 |||||  
753 |||||  
754 |||||  
755 |||||  
756 |||||  
757 |||||  
758 |||||  
759 |||||  
760 |||||  
761 |||||  
762 |||||  
763 |||||  
764 |||||  
765 |||||  
766 |||||  
767 |||||  
768 |||||  
769 |||||  
770 |||||  
771 |||||  
772 |||||  
773 |||||  
774 |||||  
775 |||||  
776 |||||  
777 |||||  
778 |||||  
779 |||||  
780 |||||  
781 |||||  
782 |||||  
783 |||||  
784 |||||  
785 |||||  
786 |||||  
787 |||||  
788 |||||  
789 |||||  
790 |||||  
791 |||||  
792 |||||  
793 |||||  
794 |||||  
795 |||||  
796 |||||  
797 |||||  
798 |||||  
799 |||||  
800 |||||  
801 |||||  
802 |||||  
803 |||||  
804 |||||  
805 |||||  
806 |||||  
807 |||||  
808 |||||  
809 |||||  
810 |||||  
811 |||||  
812 |||||  
813 |||||  
814 |||||  
815 |||||  
816 |||||  
817 |||||  
818 |||||  
819 |||||  
820 |||||  
821 |||||  
822 |||||  
823 |||||  
824 |||||  
825 |||||  
826 |||||  
827 |||||  
828 |||||  
829 |||||  
830 |||||  
831 |||||  
832 |||||  
833 |||||  
834 |||||  
835 |||||  
836 |||||  
837 |||||  
838 |||||  
839 |||||  
840 |||||  
841 |||||  
842 |||||  
843 |||||  
844 |||||  
845 |||||  
846 |||||  
847 |||||  
848 |||||  
849 |||||  
850 |||||  
851 |||||  
852 |||||  
853 |||||  
854 |||||  
855 |||||  
856 |||||  
857 |||||  
858 |||||  
859 |||||  
860 |||||  
861 |||||  
862 |||||  
863 |||||  
864 |||||  
865 |||||  
866 |||||  
867 |||||  
868 |||||  
869 |||||  
870 |||||  
871 |||||  
872 |||||  
873 |||||  
874 |||||  
875 |||||  
876 |||||  
877 |||||  
878 |||||  
879 |||||  
880 |||||  
881 |||||  
882 |||||  
883 |||||  
884 |||||  
885 |||||  
886 |||||  
887 |||||  
888 |||||  
889 |||||  
890 |||||  
891 |||||  
892 |||||  
893 |||||  
894 |||||  
895 |||||  
896 |||||  
897 |||||  
898 |||||  
899 |||||  
900 |||||  
901 |||||  
902 |||||  
903 |||||  
904 |||||  
905 |||||  
906 |||||  
907 |||||  
908 |||||  
909 |||||  
910 |||||  
911 |||||  
912 |||||  
913 |||||  
914 |||||  
915 |||||  
916 |||||  
917 |||||  
918 |||||  
919 |||||  
920 |||||  
921 |||||  
922 |||||  
923 |||||  
924 |||||  
925 |||||  
926 |||||  
927 |||||  
928 |||||  
929 |||||  
930 |||||  
931 |||||  
932 |||||  
933 |||||  
934 |||||  
935 |||||  
936 |||||  
937 |||||  
938 |||||  
939 |||||  
940 |||||  
941 |||||  
942 |||||  
943 |||||  
944 |||||  
945 |||||  
946 |||||  
947 |||||  
948 |||||  
949 |||||  
950 |||||  
951 |||||  
952 |||||  
953 |||||  
954 |||||  
955 |||||  
956 |||||  
957 |||||  
958 |||||  
959 |||||  
960 |||||  
961 |||||  
962 |||||  
963 |||||  
964 |||||  
965 |||||  
966 |||||  
967 |||||  
968 |||||  
969 |||||  
970 |||||  
971 |||||  
972 |||||  
973 |||||  
974 |||||  
975 |||||  
976 |||||  
977 |||||  
978 |||||  
979 |||||  
980 |||||  
981 |||||  
982 |||||  
983 |||||  
984 |||||  
985 |||||  
986 |||||  
987 |||||  
988 |||||  
989 |||||  
990 |||||  
991 |||||  
992 |||||  
993 |||||  
994 |||||  
995 |||||  
996 |||||  
997 |||||  
998 |||||  
999 |||||  
1000 |||||

RESULT 7  
AC005490/c  
LOCUS Homo sapiens UWGC:gl397a051 from 7p14-15, complete sequence.  
DEFINITION AC005490  
ACCESSION AC005490  
VERSION AC005490.1 GI:3420065  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 41339)  
AUTHORS Iadonato,S.P., Yu,J., Wong,G.K.-S., Magness,C.L., Green,E.D., Green,P. and Olson,M.V.  
TITLE Large-scale MCD Mapping and Sequencing of Human Chromosome 7  
JOURNAL Unpublished (1998)  
REFERENCE 2 (bases 1 to 41339)  
AUTHORS Iadonato,S.P., Yu,J., Wong,G.K.-S., Magness,C.L., Green,E.D., Green,P. and Olson,M.V.  
TITLE Direct Submission  
JOURNAL Submitted (15-AUG-1998) Human Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA  
REMARK University of Washington Human Genome Center  
Box 352145 Seattle, WA 98195  
Contact: Shawn Iadonato (iadonato@u.washington.edu)  
Overlapping Sequences:  
5': UWGC:gl211a139 (Genbank Accession: AC004397)  
3': UWGC:gl397a211

Sequence Quality Assessment:  
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

Double stranded (DS) coverage: 84.1%  
DS or two chemistry coverage: 99.3%  
Single stranded regions: 1  
-----  
Sequence Validation:

108 GCTAGTATCACTGAAACACAGCAGCAATGGTGTCTCGATGCACAGAACCATCT 167  
109 |||||  
110 |||||  
111 |||||  
112 |||||  
113 |||||  
114 |||||  
115 |||||  
116 |||||  
117 |||||  
118 |||||  
119 |||||  
120 |||||  
121 |||||  
122 |||||  
123 |||||  
124 |||||  
125 |||||  
126 |||||  
127 |||||  
128 |||||  
129 |||||  
130 |||||  
131 |||||  
132 |||||  
133 |||||  
134 |||||  
135 |||||  
136 |||||  
137 |||||  
138 |||||  
139 |||||  
140 |||||  
141 |||||  
142 |||||  
143 |||||  
144 |||||  
145 |||||  
146 |||||  
147 |||||  
148 |||||  
149 |||||  
150 |||||  
151 |||||  
152 |||||  
153 |||||  
154 |||||  
155 |||||  
156 |||||  
157 |||||  
158 |||||  
159 |||||  
160 |||||  
161 |||||  
162 |||||  
163 |||||  
164 |||||  
165 |||||  
166 |||||  
167 |||||  
168 |||||  
169 |||||  
170 |||||  
171 |||||  
172 |||||  
173 |||||  
174 |||||  
175 |||||  
176 |||||  
177 |||||  
178 |||||  
179 |||||  
180 |||||  
181 |||||  
182 |||||  
183 |||||  
184 |||||  
185 |||||  
186 |||||  
187 |||||  
188 |||||  
189 |||||  
190 |||||  
191 |||||  
192 |||||  
193 |||||  
194 |||||  
195 |||||  
196 |||||  
197 |||||  
198 |||||  
199 |||||  
200 |||||  
201 |||||  
202 |||||  
203 |||||  
204 |||||  
205 |||||  
206 |||||  
207 |||||  
208 |||||  
209 |||||  
210 |||||  
211 |||||  
212 |||||  
213 |||||  
214 |||||  
215 |||||  
216 |||||  
217 |||||  
218 |||||  
219 |||||  
220 |||||  
221 |||||  
222 |||||  
223 |||||  
224 |||||  
225 |||||  
226 |||||  
227 |||||  
228 |||||  
229 |||||  
230 |||||  
231 |||||  
232 |||||  
233 |||||  
234 |||||  
235 |||||  
236 |||||  
237 |||||  
238 |||||  
239 |||||  
240 |||||  
241 |||||  
242 |||||  
243 |||||  
244 |||||  
245 |||||  
246 |||||  
247 |||||  
248 |||||  
249 |||||  
250 |||||  
251 |||||  
252 |||||  
253 |||||  
254 |||||  
255 |||||  
256 |||||  
257 |||||  
258 |||||  
259 |||||  
260 |||||  
261 |||||  
262 |||||  
263 |||||  
264 |||||  
265 |||||  
266 |||||  
267 |||||  
268 |||||  
269 |||||  
270 |||||  
271 |||||  
272 |||||  
273 |||||  
274 |||||  
275 |||||  
276 |||||  
277 |||||  
278 |||||  
279 |||||  
280 |||||  
281 |||||  
282 |||||  
283 |||||  
284 |||||  
285 |||||  
286 |||||  
287 |||||  
288 |||||  
289 |||||  
290 |||||  
291 |||||  
292 |||||  
293 |||||  
294 |||||  
295 |||||  
296 |||||  
297 |||||  
298 |||||  
299 |||||  
300 |||||  
301 |||||  
302 |||||  
303 |||||  
304 |||||  
305 |||||  
306 |||||  
307 |||||  
308 |||||  
309 |||||  
310 |||||  
311 |||||  
312 |||||  
313 |||||  
314 |||||  
315 |||||  
316 |||||  
317 |||||  
318 |||||  
319 |||||  
320 |||||  
321 |||||  
322 |||||  
323 |||||  
324 |||||  
325 |||||  
326 |||||  
327 |||||  
328 |||||  
329 |||||  
330 |||||  
331 |||||  
332 |||||  
333 |||||  
334 |||||  
335 |||||  
336 |||||  
337 |||||  
338 |||||  
339 |||||  
340 |||||  
341 |||||  
342 |||||  
343 |||||  
344 |||||  
345 |||||  
346 |||||  
347 |||||  
348 |||||  
349 |||||  
350 |||||  
351 |||||  
352 |||||  
353 |||||  
354 |||||  
355 |||||  
356 |||||  
357 |||||  
358 |||||  
359 |||||  
360 |||||  
361 |||||  
362 |||||  
363 |||||  
364 |||||  
365 |||||  
366 |||||  
367 |||||  
368 |||||  
369 |||||  
370 |||||  
371 |||||  
372 |||||  
373 |||||  
374 |||||  
375 |||||  
376 |||||  
377 |||||  
378 |||||  
379 |||||  
380 |||||  
381 |||||  
382 |||||  
383 |||||  
384 |||||  
385 |||||  
386 |||||  
387 |||||  
388 |||||  
389 |||||  
390 |||||  
391 |||||  
392 |||||  
393 |||||  
394 |||||  
395 |||||  
396 |||||  
397 |||||  
398 |||||  
399 |||||  
400 |||||  
401 |||||  
402 |||||  
403 |||||  
404 |||||  
405 |||||  
406 |||||  
407 |||||  
408 |||||  
409 |||||  
410 |||||  
411 |||||  
412 |||||  
413 |||||  
414 |||||  
415 |||||  
416 |||||  
417 |||||  
418 |||||  
419 |||||  
420 |||||  
421 |||||  
422 |||||  
423 |||||  
424 |||||  
425 |||||  
426 |||||  
427 |||||  
428 |||||  
429 |||||  
430 |||||  
431 |||||  
432 |||||  
433 |||||  
434 |||||  
435 |||||  
436 |||||  
437 |||||  
438 |||||  
439 |||||  
440 |||||  
441 |||||  
442 |||||  
443 |||||  
444 |||||  
445 |||||  
446 |||||  
447 |||||  
448 |||||  
449 |||||  
450 |||||  
451 |||||  
452 |||||  
453 |||||  
454 |||||  
455 |||||  
456 |||||  
457 |||||  
458 |||||  
459 |||||  
460 |||||  
461 |||||  
462 |||||  
463 |||||  
464 |||||  
465 |||||  
466 |||||  
467 |||||  
468 |||||  
469 |||||  
470 |||||  
471 |||||  
472 |||||  
473 |||||  
474 |||||  
475 |||||  
476 |||||  
477 |||||  
478 |||||  
479 |||||  
480 |||||  
481 |||||  
482 |||||  
483 |||||  
484 |||||  
485 |||||  
486 |||||  
487 |||||  
488 |||||  
489 |||||  
490 |||||  
491 |||||  
492 ||

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
 Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,  
 Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,  
 Collamore,A., Cook,A., Cooke,P., Corum,B., DeArelano,K.,  
 Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,  
 Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,  
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,  
 Hall,J., Horton,D., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kanat,A., Karatas,A., Kellis,C., Landers,T., Levine,R.,  
 Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,  
 MacDonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,  
 Meldrim,J., Meneus,L., Mihova,T., Mianga,V., Murphy,T., Naylor,J.,  
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,  
 O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
 Rachupka,A., Ranasingh,U., Raymond,C., Ratta,R., Rise,C., Rogov,P.,  
 Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,  
 Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,  
 Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,  
 Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,  
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (10-MAR-2003) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Mar 10, 2003 this sequence version replaced gi:28369510.  
 All repeats were identified using RepeatMasker:  
 Smith, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIGR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence.submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L22768  
 Center clone name: 437\_I.18  
 ----- Summary Statistics  
 Sequencing vector: Plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 190850 bases at least Q40  
 Consensus quality: 191537 bases at least Q30  
 Consensus quality: 191930 bases at least Q20  
 Insert size: 183000; agarose-fp  
 Insert size: 192272; sum-of-contigs  
 Quality coverage: 11.0 in Q20 bases; agarose-fp  
 Quality coverage: 10.5 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 5 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence.  
 \* As soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 \* 1  
 \* 23120: contig of 23120 bp in length  
 \* 23121 23220: gap of 100 bp  
 \* 23221 30232: contig of 7012 bp in length  
 \* 30232 30333: gap of 100 bp  
 \* 30333 53334: contig of 23002 bp in length  
 \* 53334 53434: gap of 100 bp  
 \* 53435 141535: contig of 88101 bp in length  
 \* 141535 141635: gap of 100 bp  
 \* 141635 192672: contig of 51037 bp in length.  
 \* Location/Qualifiers  
 \* 1..192672  
 \* /organism="Mus musculus"  
 \* /mol\_type="genomic DNA"  
 \* /db\_xref="taxon:10090"  
 \* /clone="RP23-437118"  
 \* /clone\_lib="RPCI-23 Female Mouse BAC"  
 \* 1..23120  
 \* /note="assembly fragment  
 ature

```

clone_end:SP6
vector_side:left
23221..30232
/notes="assembly_fragment"
30333..53334
/notes="assembly_fragment"
53435..141535
/notes="assembly_fragment"
141636..192672
/notes="assembly_fragment"
clone_end:T7
vector_side:right"

FEATURES
    source
        1..221190
            /organism="Danio rerio"
            /mol_type="genomic DNA"
            /db_xref="taxon:7955"
            /clone="DKEY-33M14"
            /clone_lib="DanioKey"
        1..4748
            /note="assembly_fragment:00892"
            fragment_chain:1"
        4849..26640
            /note="assembly_fragment:01375"
            fragment_chain:1"
        26741..39754
            /note="assembly_fragment:01249"
            fragment_chain:1"
        39855..79214
            /note="assembly_fragment:00365"
            fragment_chain:1"
        79315..193130
            /note="assembly_fragment:00635"
            fragment_chain:1"
        193231..204142
            /note="assembly_fragment:00668"
            fragment_chain:1"
        204243..207209
            /note="assembly_fragment:00870.0"
            fragment_chain:1"
        207310..221190
            /note="assembly_fragment:00990"
            fragment_chain:1"

ORIGIN
Query Match      4.4%; Score 21; DB 2; Length 192672;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 AATTAATCAAGGAGAGAGAAAT 64
Db 21707 AATTAATCAAGGAGAGAGAAAT 21727

RESULT 9
BX571704/c      221190 bp DNA linear HTG 24-SEP-2003
LOCUS           221190 bp DNA linear HTG 24-SEP-2003
DEFINITION      Danio rerio clone DKEY-33M14, WORKING DRAFT SEQUENCE, 8 unordered
                pieces.
ACCESSION       BX571704
VERSION         BX571704.3 GI:35209958
KEYWORDS        HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE          Danio rerio (zebrafish)
ORGANISM        Danio rerio
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
                Cypriniformes; Cyprinidae; Danio.
REFERENCE       1 (bases 1 to 221190)
                McLaren S.
                Direct Submission
                Submitted (23-SEP-2003) Wellcome Trust Sanger Institute, Hinxton,
                Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
                zf1sh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
                On Sep 24, 2003 this sequence version replaced gi:32879563.
                ----- Genome Center
                Center: Wellcome Trust Sanger Institute
                Center code: SC
                Web site: http://www.sanger.ac.uk
                Contact: zf1sh-help@sanger.ac.uk
                ----- Project Information
                Center project name: zk33M14
                ----- Summary Statistics
                Assembly program: XGAP4; version 4.5
                Chemistry: Dye-terminator; 100% of reads
                Consensus quality: 218263 bases at least Q40
                Consensus quality: 218835 bases at least Q30
                Consensus quality: 219365 bases at least Q20
                Insert size: 220490; sum-of-contigs
                Insert size: 180894; 7.9% error; agarose-fp
                Quality coverage: 6.34x in Q20 bases; sum-of-contigs Quality
                coverage: 7.77x in Q20 bases; agarose-fp
                -----
                * NOTE: This is a 'working draft' sequence. It currently
                * consists of 8 contigs. The true order of the pieces
                * is not known and their order in this sequence record is
                * arbitrary. Gaps between the contigs are represented as
                * runs of N, but the exact sizes of the gaps are unknown.
                * This record will be updated with the finished sequence
                * as soon as it is available and the accession number will
                * be preserved.
                *
                1 4748: contig of 4748 bp in length
                * 4749 4848: gap of 100 bp
                * 4849 26640: contig of 21792 bp in length
                * 26641 26740: gap of 100 bp
                *

```

Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN:  
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,  
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,  
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Oka, Y.,  
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Yoshino, M., and Hayashizaki, Y.  
Collection, mapping, and annotation of over 28,000 cDNA clones from  
japonica rice  
Science 301 (5631), 376-379 (2003)  
22752273  
12869764  
2 (bases 1 to 1046)  
ADACHI, J., AIZAWA, K., AKIMURA, T., ARAKAWA, T., CARNINCI, P., DOI, K.,  
FUJIMURA, T., FUKUDA, S., HANAGAKI, T., HARA, A., HASHIZUME, W.,  
HAYASHIDA, K., HAYASHIZAKI, Y., HAYATSU, N., HIRAMOTO, K., HIROAKA, T.,  
HORI, F., HOTTA, I., IIDA, J., IIDA, Y., IKEDA, R., IMAMURA, K.,  
IMOTANI, K., ISHIBIKI, J., ISHII, Y., ISHIKAWA, M., ITOH, M., KAGAWA, I.,  
KANAGAWA, S., KATOH, H., KAWAGASHIRA, N., KAWAI, J., KAWAMATA, M.,  
KIKUCHI, S., KISHIKAWA-HIROZANE, T., KISHIMOTO, N., KOBAYASHI, M.,  
KODAMA, T., KOJIMA, K., KOJIMA, Y., KONDO, S., KONNO, H., KOUDA, M.,  
KOYA, S., KURIHARA, C., KUROSAKI, T., KUSUMEGI, T., LI, C., LU, M.,  
MASUDA, H., MATSUBARA, K., MATSUYAMA, T., MIURA, J., MIYAZAKI, A.,  
MIZUNO, K., MURAKAMI, K., MURATA, M., NAGATA, T., NAKAMURA, M.,  
NAMIKI, T., NARIKAWA, R., NIIKURA, J., NISHI, K., NOMURA, K.,  
NUMASAKI, R., OHNEDA, E., OHNO, M., OHTSUKI, K., OKA, M., OKA, H.,  
OSATO, N., OTA, Y., OTOMO, Y., RYU, R., SAITOH, H., SAKAI, C., SAKAI, K.,  
SAKAZUME, N., SANO, H., SASAKI, D., SATO, K., SATOH, K., SHIBATA, K.,  
SHINAGAWA, A., SHIRAKI, T., SHISHIKI, T., SOGABE, Y., SUGANO, S.,  
SUGIYAMA, A., SUZUKI, K., SUZUKI, Y., TAGAMI, M., TAGAMI-TAKEDA, Y.,  
TAGAWA, T., TAKAHASHI, F., TAKAKU-AKASHIRA, S., TANAKA, T., TOMARU, A.,  
TOYA, T., TSUNODA, Y., UEDA, M., WAKI, K., XIE, Q., YAHAGI, W.,  
YAMADA, H., YANAMOTO, M., YASUNISHI, A., YAZAKI, J., YOKOMIZO, S., and  
YOSHIMURA, A.  
Direct Submission  
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of  
Agrobiological Sciences, Department of Molecular Genetics, Head of  
Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki  
305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,  
Tel: 81-29-838-7007, Fax: 81-29-838-7007)  
This clone is one of the 28K full-length cDNA clones from japonica  
rice.  
URL: <http://cdna01.dna.affrc.go.jp/cdna/>  
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,  
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,  
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,  
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and  
Yamamoto, M.  
FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,  
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,  
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,  
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Masuda, H., Murata, S.,  
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,  
Yoshimura, A., Matsubara, K., and Murakami, K.  
Genome Exploration Research Group in Riken Genomic Sciences Center  
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,  
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,  
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,  
Hiroaka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,  
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,  
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,  
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,  
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,  
Takaku-Akashira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,  
Yasunishi, A., and Hayashizaki, Y.  
Location/Qualifiers  
1..1046  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultiVar="Nipponbare"  
/db\_xref="taxon:39947"

ORIGIN  
Query Match 4.2%; Score 20; DB 8; Length 1046;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
/clone="006-211-B02"  
QY 413 GAGCTTGATTTTCATTGGA 432  
|||||  
DB 465 GAGCTTGATTTTCATTGGA 484  
|||||  
RESULT 11  
SCYKL041W  
LOCUS SCYKL041W 1490 bp DNA linear PLN 18-MAR-1998  
DEFINITION S.cerevisiae chromosome XI reading frame ORF YKL041w.  
ACCESSION Z28041 Y13137  
VERSION Z28041.1 GI:486051  
KEYWORDS Saccharomyces cerevisiae (baker's yeast)  
SOURCE Saccharomyces cerevisiae  
ORGANISM Saccharomyces cerevisiae  
REFERENCE 1 (bases 1 to 1490)  
AUTHORS Furnelle, B., Skala, J., van Dyck, L., Tettelin, H. and Goffeau, A.  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1490)  
AUTHORS MIPS  
TITLE Direct Submission  
JOURNAL Submitted (09-MAY-1994) Data collected by MIPS on behalf of the  
European yeast chromosome XI sequencing project. MIPS at the  
Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152  
Martinsried, FRG; E-mail: Mewes@ehpmic.mips.biochem.mpg.de  
FEATURES  
Location/Qualifiers  
1..1490  
/organism="Saccharomyces cerevisiae"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4932"  
/chromosome="XI"  
576..1250  
/gene="VPS24"  
576..1250  
/gene="VPS24"  
/note="ORF YKL041w"  
/codon\_start=1  
/protein\_id="CAA81876.1"  
/db\_xref="GI:486052"  
/db\_xref="GOA:P36095"  
/db\_xref="SGD:S0001524"  
/db\_xref="SWISS-PROT:P36095"  
/translation="MDYIKKAIWGDDPKQQRRTSVLRKNGRIEKLRLTVLQNK  
TQOLIKAKKNDVRLVYAKELYQINKQDMYTSRAQDLSVRMKIDEAIRNLTLS  
NQMDASAGLMREVNLRPLQRLNTMIELKELKSGIISENVDDTMSVGDVGDEMD  
EAVDVEVNKIVEQYTNKRFKNVDQVPTVELAANESEQIPDEKVDDEADRMVNER  
LRALQN"  
ORIGIN  
Query Match 4.2%; Score 20; DB 8; Length 1490;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 416 CTGTGATTTTCATTGAAGG 435  
|||||  
DB 1462 CTGTGATTTTCATTGAAGG 1481  
|||||  
RESULT 12  
AX008583  
LOCUS AX008583 1557 bp DNA linear PAT 06-SEP-2000  
DEFINITION Sequence 1 from Patent WO9966056.  
ACCESSION AX008583  
VERSION AX008583.1 GI:9996133  
KEYWORDS

```

SOURCE      Chlamydomonas reinhardtii
ORGANISM    Chlamydomonas reinhardtii
            Chlamydomonadaceae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonas.

REFERENCE   1
AUTHORS    Ball,S.
TITLE      Method for obtaining modified polysaccharides
JOURNAL    Patent: WO 996056-A 1 23-DEC-1999;
            BALL STEVEN (FR); BIOENNA (FR)
FEATURES   Location/Qualifiers
            source
              1..1557
              /organism="Chlamydomonas reinhardtii"
              /mol_type="unassigned DNA"
              /db_xref="taxon:3055"

ORIGIN
Query Match      4.2%; Score 20; DB 6; Length 1557;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      87 AGCCCGCTTTACAAAGTGCC 106
Db      1441 AGCCCGCTTTACAAAGTGCC 1460

RESULT 13
LOCUS    SCYKL040C
DEFINITION S.cerevisiae chromosome XI reading frame ORF YKL040c.
ACCESSION Z28040 Y13137
VERSION   Z28040.1 GI:486049
KEYWORDS  Saccharomyces cerevisiae (baker's yeast)
SOURCE    Saccharomyces cerevisiae
ORGANISM  Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

REFERENCE   1
AUTHORS    Purnelle,B., Skala,J., van Dyck,L., Tettelin,H. and Goffeau,A.
JOURNAL    Unpublished
TITLE      (bases 1 to 1804)
AUTHORS    MIPS.
JOURNAL    Submitted (09-MAY-1994) Data collected by MIPS on behalf of the
            European yeast chromosome XI sequencing project. MIPS at the
            Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152
            Martinsried, FRG; E-mail: Mewese@hgm.mc.mpg.de
FEATURES   Location/Qualifiers
            source
              1..1804
              /organism="Saccharomyces cerevisiae"
              /mol_type="genomic DNA"
              /db_xref="taxon:4932"
              /chromosome="XI"
              complement(241..1011)
              /note="unnamed protein product; ORF YKL040c"
              /codon_start=1
              /protein_id="CAA81875.1"
              /db_xref="GI:486050"
              /db_xref="GOA:P32860"
              /db_xref="SWISS-PROT:P32860"
              /translation="MFKSVAKLGKSPFYLNRSORLIHKTLTPFNENALKFLSTDGEM
            LQTRGSKIYKNTDENLHNSKLAQQIFLQCPGVBSLMIGDDFLTNDRMVHNSI
            LQTRGSKIYKNTDENLHNSKLAQQIFLQCPGVBSLMIGDDFLTNDRMVHNSI
            IDTRIPALIEDGGDIDYRGWDFKGTGYVYLRLOGACTSCSSSEVTLYKYGIESMLKHVY
            DEKVEIQIWDPEQIALKFDFLEKLESKNTSHEK"

CDS
            1..1804
            /organism="Saccharomyces cerevisiae"
            /mol_type="genomic DNA"
            /db_xref="taxon:4932"
            /chromosome="XI"
            complement(241..1011)
            /note="unnamed protein product; ORF YKL040c"
            /codon_start=1
            /protein_id="CAA81875.1"
            /db_xref="GI:486050"
            /db_xref="GOA:P32860"
            /db_xref="SWISS-PROT:P32860"
            /translation="MFKSVAKLGKSPFYLNRSORLIHKTLTPFNENALKFLSTDGEM
            LQTRGSKIYKNTDENLHNSKLAQQIFLQCPGVBSLMIGDDFLTNDRMVHNSI
            LQTRGSKIYKNTDENLHNSKLAQQIFLQCPGVBSLMIGDDFLTNDRMVHNSI
            IDTRIPALIEDGGDIDYRGWDFKGTGYVYLRLOGACTSCSSSEVTLYKYGIESMLKHVY
            DEKVEIQIWDPEQIALKFDFLEKLESKNTSHEK"

ORIGIN
Query Match      4.2%; Score 20; DB 8; Length 1804;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      416 CTTGTAATTTTCATTGAAGG 435
            |||||||||||||||||||

```

```

Db      212 CTTGTAATTTTCATTGAAGG 231

RESULT 14
LOCUS    AF307842
DEFINITION Chlamydomonas reinhardtii 4-alpha-glucanotransferase (STAl1) mRNA,
            complete cds.
ACCESSION AF307842
VERSION   AF307842.1 GI:11095334
KEYWORDS  Chlamydomonas reinhardtii
            Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1
AUTHORS    Wasteson,F., Rai,J.P., Dauvillee,D., Myers,A.M., James,M.G.,
            Schlitching,R., Giersch,C., Ball,S.G. and D'Huist,C.
TITLE      STAl1, a Chlamydomonas reinhardtii Locus Required for Normal Starch
            Granule Biogenesis, Encodes Disproportionating Enzyme. Further
            Evidence for a Function of alpha-1,4 Glucanotransferases during
            Starch Granule Biosynthesis in Green Algae
JOURNAL    Plant Physiol. 132 (1), 137-145 (2003)
PUBMED    12746519
REFERENCE   2
AUTHORS    Wasteson,F., Rai,J.P., Dauvillee,D., Myers,A.M. and Ball,S.G.
TITLE      Direct Submission
JOURNAL    Submitted (22-SEP-2000) UMR8576 of CNRS, Laboratory of Biological
            Chemistry, University of Sciences and Technologies of Lille,
            Villeneuve d'Ascq Cedex 59655, France
FEATURES   Location/Qualifiers
            source
              1..2204
              /organism="Chlamydomonas reinhardtii"
              /mol_type="mRNA"
              /db_xref="taxon:3055"
              1..2204
              /gene="STAl1"
              130..1887
              /gene="STAl1"
              /note="disproportionating enzyme; D-enzyme"
              /citation=[1]
              /codon_start=1
              /product="4-alpha-glucanotransferase"
              /protein_id="AAC39839.1"
              /db_xref="GI:11095335"
              /translation="WIPLRAPQGVRYTQIDRSQVLPQLRSRPAQSVSHVAAAGSRRRA
            VSRSVRAFAMGTPTKTTGAPGVGLGKDYDQKIPSYNKRERAGILLHTSLPGPYGIG
            DIGEAKRFVDMLADHGMQWQLPLVPPDPMTYSPYSGTDANGCNFLVSIIEELIKD
            GLBESETPRVPIADVDYPAVAAKLLPLKRAAQRLLKEDRFLRREYLYKREKHP
            WEDSALFDVARNLPESLQAWMPSPRLRQXKALKKEFEETNKDAIDRFVVIQYFF
            EKQWKAIRSYANGKIGKILGDMPIYVGSHSADVWNSHLELNEAGLPEQYSGVPPDA
            FSATGQLWGSPLYKWPAAHKKGFGKWTARTARVARTLELDECRIDHFRFAGYWSVDANE
            TTAMNGNRQFGLELFTAMKALGAVPIAEIDLGIVITDVVALREAGAPGMVVLQF
            AWGGPGNVHLPHNHYENCFYPTGTHDNETAVGWRFGSANDTDSYI:KSLRTDGGDI
            AWDFTACMAAAPTCTVIMQDMVRLONTARMTPTGTAAGNWRMRMGDGNVWNSLKKE
            GEDLRKVAHDTNRLPKPKA"

ORIGIN
Query Match      4.2%; Score 20; DB 8; Length 2204;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      87 AGCCCGCTTTACAAAGTGCC 106
            |||||||||||||||||||
Db      1159 AGCCCGCTTTACAAAGTGCC 1178

RESULT 15
LOCUS    BC051144/c
DEFINITION Mus musculus RIKEN cDNA C330007P06 gene, mRNA (cDNA clone MGC:58463
            IMAGE:658980), complete cds.
ACCESSION BC051144

```

BC051144.1 GI:30047895  
MGC.  
Mus musculus (house mouse)  
ORGANISM  
REFERENCE  
AUTHORS  
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
Altshuler,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
Diatchenko,L., Marusina,K., Farmer,A., Rubin,G.M., Hong,L.,  
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
Schectt,R.B., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,  
Carninci,P., Prange,C., Raha,S., Lequellano,N.A., Peters,G.J.,  
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,  
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,  
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
Fahey,J., Helton,E., Kettner,M., Madan,A., Rodriguez,S.,  
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
Butterfield,Y.S., Krzyzinski,M.I., Skalska,U., Smailus,D.E.,  
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
22388257  
12477932  
2 (bases 1 to 2709)  
Strausberg,R.  
Direct Submission  
Submitted (11-APR-2003) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Dr. Michael Brownstein  
cDNA Library Preparation: Michael Brownstein / Ted Usdin  
Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
Leticia Hsiao, Martin Krzyzinski, Reta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline  
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,  
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 45 Row: e Column: 3  
This clone was selected for full length sequencing because it  
passed the following selection criteria: GenomeScan gene  
prediction.  
Location/Qualifiers  
1. .2709  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="MGC:58463 IMAGE:6588980"  
/tissue\_type="Brain, mouse"  
/clone\_lib="NIH\_MGC\_144"  
/lab\_host="DH10B"  
/note="vector: pDNR-LIB"  
1. .2709  
gene

ORIGIN  
Query Match 4.2%; Score 20; DB 10; Length 2709;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 369 AATTATGTTTCATTATTATT 388  
Db 2398 AATTATGTTTCATTATTATT 2379  
|||||  
Search completed: July 31, 2004, 10:16:38  
Job time : 2310 secs

/genes="C330007P06Rik"  
/db\_xref="LocusID:77644"  
/db\_xref="MGI:1924894"  
28..696  
/codon\_start=1  
/product="C330007P06Rik protein"  
/protein\_id="AAH51144.1"  
/db\_xref="GI:30047896"  
/db\_xref="LocusID:77644"  
/translation="MPKVSRVSVCSGSDTDRREYDDGKPLHVVYCLGQMVVLDCQ  
LEKLPWRPDRSRVIDAARAHKFCNTDEETTYLRRPGLEROYRKKCAKGLPLFY  
QSQPKNAPVTFTVDGAVYKFGCGFKNTIYTKQBPFPKKVMTKTKDKMGKSSVTVS  
TIDEEERIEAREVADSYAQNKVLEKQLERKMGSKRRRLQELAELEAKKAKMKGTLLD  
NQFK"

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 06:52:19 ; Search time 307 Seconds  
(without alignments)  
6586.783 Million cell updates/sec

Title: US-10-020-540A-1

Perfect score: 476

Sequence: 1 aactatgctgacgaagat.....aacctccctgctacaacc 476

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001as.\*
- 5: Geneseq2001bs.\*
- 6: Geneseq2002s.\*
- 7: Geneseq2003as.\*
- 8: Geneseq2003bs.\*
- 9: Geneseq2003cs.\*
- 10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	476	100.0	476	6 AAD39132	Aad39132 Strawberry
2	476	100.0	518	6 AAD39133	Aad39133 Strawberry
3	476	100.0	1057	6 ABX09816	Abx09816 Promoter
4	476	100.0	1071	6 ABX09826	Abx09826 SVBV type
5	369	77.5	651	6 AAD39134	Aad39134 Strawberry
6	369	77.5	701	6 AAD39135	Aad39135 Strawberry
7	37.6	7.5	3294	4 ABL26902	AbL26902 Drosophila
8	37.6	7.9	78845	3 AA81463	AA81463 N. mening
9	37.6	7.9	110000	3 AA81490	Continuation (9 of
10	37.6	7.9	349980	3 AAF21608	Aaf21608 Neisseria
11	36.8	7.7	6012	4 AAS46724	Aas46724 Tumour su
12	36.6	7.7	17211	6 ABL32655	AbL32655 Human imm
13	36.4	7.6	285	7 ABZ38820	Abz38820 N. gonorr
14	36.2	7.6	110000	7 AAD53224_2	Continuation (3 of
15	36.2	7.6	110000	7 AAD53224_3	Continuation (4 of
16	35.8	7.5	963	3 AAA70197	Aaa70197 Plasmodiu
17	35.2	7.4	160	5 ABV56034	Abv56034 Human pro
18	35.2	7.4	8992	6 ABK31421	Abk31421 Signal tr
19	35	7.4	458	4 AAS28807	Aas28807 Human imm
20	35	7.4	458	4 AAS27944	Aas27944 Novel cdn
21	35	7.4	458	4 ABA06606	AbA06606 Human cdn
22	35	7.4	458	6 ABV83943	Abv83943 Human pol
23	35	7.4	458	9 ADB31532	Adb31532 Human cdn

24	35	7.4	502	4 ABA06349	AbA06349 Human CDN
25	35	7.4	502	6 ABV83686	Abv83686 Human pol
26	35	7.4	504	4 AAS28903	Aas28903 Human imm
27	35	7.4	504	4 AAS28336	Aas28336 Genomic s
28	35	7.4	504	9 ADB31744	Adb31744 Human nov
29	34.8	7.3	112190	4 AAD44801	Aad44801 Human GPC
30	34.6	7.3	923	4 AAD05641	Aad05641 Human sec
31	34.6	7.3	923	4 AAD08435	Aad08435 Human sec
32	34.6	7.3	923	7 ADA56482	Ada56482 Gene enco
33	34.6	7.3	923	7 ADA40320	Ada40320 Human sec
34	34.6	7.3	923	9 AAD73842	Aad73842 Human sec
35	34.6	7.3	6314	4 AAK90962	Aak90962 Human dig
36	34.6	7.3	6314	5 AAS31997	Aas31997 Human liv
37	34.6	7.3	6314	6 ABN90352	Abn90352 Human liv
38	34.6	7.3	9289	4 AAK90960	Aak90960 Human dig
39	34.6	7.3	9289	5 AAS31995	Aas31995 Human liv
40	34.6	7.3	9289	6 ABN90350	Abn90350 Human liv
41	34.6	7.3	17865	4 AAK90961	Aak90961 Human dig
42	34.6	7.3	17865	5 AAS31996	Aas31996 Human liv
43	34.6	7.3	17865	6 ABN90351	Abn90351 Human liv
44	34.6	7.3	49047	8 ADA02633	Ada02633 Mouse Dnt
45	34.6	7.3	49047	9 ADB72371	Adb72371 Mouse Dnt

ALIGNMENTS

RESULT 1

AD39132	AD39132	standard; DNA; 476 BP.
XX	XX	
AC	AD39132;	
XX	XX	
DT	04-OCT-2002 (first entry)	
XX	XX	
DE	Strawberry vein banding virus (SVBV) promoter fragment DNA.	
XX	XX	
KW	Strawberry vein banding virus promoter; transgenic plant; transgenic; ds.	
XX	XX	
OS	Strawberry vein banding virus.	
XX	XX	
PN	WO200240691-A2.	
XX	XX	
PD	23-MAY-2002.	
XX	XX	
PF	30-OCT-2001; 2001WO-US047964.	
XX	XX	
PR	01-NOV-2000; 2000US-0245354P.	
XX	XX	
PA	(MAXY-) MAXYGEN INC.	
XX	XX	
PI	Wu G, McBride K;	
XX	XX	
DR	WPI; 2002-479908/51.	
XX	XX	
PT	Novel isolated or recombinant nucleic acid comprising strawberry vein	
PT	banding virus promoter operably linked to a heterologous polynucleotide,	
PT	useful for expressing the heterologous polynucleotide in a plant cell.	
XX	XX	
PS	Claim 1; Page 47; 49pp; English.	
XX	XX	
CC	The invention relates to an isolated/recombinant nucleic acid comprising	
CC	a strawberry vein banding virus (SVBV) promoter operably linked to a	
CC	heterologous polynucleotide. The invention is useful for expressing a	
CC	heterologous polynucleotide in a plant cell, by introducing the nucleic	
CC	acid into a plant cell, where the plant cell is present within a plant	
CC	cell. The invention is used to introduce the isolated nucleic acid into the	
CC	cell. The invention is useful to drive gene expression in plant cells and	
CC	transgenic plants, for transforming plant cells and producing transgenic	
CC	plants, for compensating for missing or altered gene expression in a	
CC	plant, for expressing an endogenous protein at higher than normal levels,	
CC	for expressing a novel gene in a plant, and to suppress expression of	
CC	endogenous plant genes. The present sequence is SVBV-E3 promoter fragment	

```
CC DNA
XX
SQ Sequence 476 BP; 164 A; 109 C; 87 G; 116 T; 0 U; 0 Other;

Query Match      100.0%; Score 476; DB 6; Length 476;
Best Local Similarity 100.0%; Pred. No. 1.8e-123;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTATGCTGATGACAGCAATATCTTAATAGCAATATTTCAGAAATTAATCAAGGAGAAA 60
   |||||
DB 1 AACTATGCTGATGACAGCAATATCTTAATAGCAATATTTCAGAAATTAATCAAGGAGAAA 60

QY 61 GAATTAATAAATCTTTTCAGATATGAAGCCGCTTTTCAAGTGGCCAGCTAGCTATCACT 120
   |||||
DB 61 GAATTAATAAATCTTTTCAGATATGAAGCCGCTTTTCAAGTGGCCAGCTAGCTATCACT 120

QY 121 GAAAAGACAGCAAGCAATGCTGCTCGATGCACAGAACACACATCTTTGCGAGCATGT 180
   |||||
DB 121 GAAAAGACAGCAAGCAATGCTGCTCGATGCACAGAACACACATCTTTGCGAGCATGT 180

QY 181 GAAGCAGCAGAGTGGTCCCAAGACGCACTCAGAAAAGGCATCTTCTACCGACACAGAA 240
   |||||
DB 181 GAAGCAGCAGAGTGGTCCCAAGACGCACTCAGAAAAGGCATCTTCTACCGACACAGAA 240

QY 241 AAAGACACACACAGCTCATCTCAACATGTAGACTGCTGTTATGCGCTGGAAGATA 300
   |||||
DB 241 AAAGACACACACAGCTCATCTCAACATGTAGACTGCTGTTATGCGCTGGAAGATA 300

QY 301 AGACTGACCCAGGCGCAGCACTAAAGAAAGAAATTAATGCAAGTGGTCTTAGCTCCACTTTA 360
   |||||
DB 301 AGACTGACCCAGGCGCAGCACTAAAGAAAGAAATTAATGCAAGTGGTCTTAGCTCCACTTTA 360

QY 361 GCTTTAATAATATGTTTCATATATATCTCTGCTTTTCTCTATATAAAGAGCTTGT 420
   |||||
DB 361 GCTTTAATAATATGTTTCATATATATCTCTGCTTTTCTCTATATAAAGAGCTTGT 420

QY 421 ATTTTCATTTGAAGCAGAGCGGCAACACACACAGAACCTCCCTGCTTACAAACC 476
   |||||
DB 421 ATTTTCATTTGAAGCAGAGCGGCAACACACACAGAACCTCCCTGCTTACAAACC 476

RESULT 2
AAD39133
ID AAD39133 standard; DNA; 518 BP.
XX
AC AAD39133;
XX
DT 04-OCT-2002 (first entry)
XX
DE Strawberry vein banding virus (SVBV) promoter DNA.
XX
KW Strawberry vein banding virus promoter; transgenic plant; transgenic; ds.
XX
OS Strawberry vein banding virus.
XX
FN WO200240691-A2.
XX
PD 23-MAY-2002.
XX
PF 30-OCT-2001; 2001WO-US047964.
XX
PR 01-NOV-2000; 2000US-0245354P.
XX
PA (MAXY-) MAXYGEN INC.
XX
PI Wu G, McBride K;
XX
DR WPI; 2002-479908/51.
XX
PT Novel isolated or recombinant nucleic acid comprising strawberry vein
PT banding virus promoter operably linked to a heterologous polynucleotide,
PT useful for expressing the heterologous polynucleotide in a plant cell.
```

```
PS Claim 1; Page 47; 48pp; English.
XX
CC The invention relates to an isolated/recombinant nucleic acid comprising
CC a strawberry vein banding virus (SVBV) promoter operably linked to a
CC heterologous polynucleotide. The invention is useful for expressing a
CC heterologous polynucleotide in a plant cell, by introducing the nucleic
CC acid into a plant cell, where the plant cell is present within a plant
CC and Agrobacterium is used to introduce the isolated nucleic acid into the
CC cell. The invention is useful to drive gene expression in plant cells and
CC transgenic plants, for transforming plant cells and producing transgenic
CC plants, for compensating for missing or altered gene expression in a
CC plant, for expressing an endogenous protein at higher than normal levels,
CC for expressing a novel gene in a plant, and to suppress expression of
CC endogenous plant genes. The present sequence is SVBV promoter DNA
XX
SQ Sequence 518 BP; 171 A; 122 C; 100 G; 125 T; 0 U; 0 Other;

Query Match      100.0%; Score 476; DB 6; Length 518;
Best Local Similarity 100.0%; Pred. No. 1.8e-123;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTATGCTGATGACAGCAATATCTTAATAGCAATATTTCAGAAATTAATCAAGGAGAAA 60
   |||||
DB 15 AACTATGCTGATGACAGCAATATCTTAATAGCAATATTTCAGAAATTAATCAAGGAGAAA 74

QY 61 GAATTAATAAATCTTTTCAGATATGAAGCCGCTTTTCAAGTGGCCAGCTAGCTATCACT 120
   |||||
DB 75 GAATTAATAAATCTTTTCAGATATGAAGCCGCTTTTCAAGTGGCCAGCTAGCTATCACT 134

QY 121 GAAAAGACAGCAAGCAATGCTGCTCGATGCACAGAACACACATCTTTGCGAGCATGT 180
   |||||
DB 135 GAAAAGACAGCAAGCAATGCTGCTCGATGCACAGAACACACATCTTTGCGAGCATGT 194

QY 181 GAAGCAGCAGAGTGGTCCCAAGACGCACTCAGAAAAGGCATCTTACCGACACAGAA 240
   |||||
DB 195 GAAGCAGCAGAGTGGTCCCAAGACGCACTCAGAAAAGGCATCTTACCGACACAGAA 254

QY 241 AAAGACACACACAGCTCATCTCAACATGTAGACTGCTGTTATGCGCTGGAAGATA 300
   |||||
DB 255 AAAGACACACACAGCTCATCTCAACATGTAGACTGCTGTTATGCGCTGGAAGATA 314

QY 301 AGACTGACCCAGGCGCAGCACTAAAGAAAGAAATTAATGCAAGTGGTCTTAGCTCCACTTTA 360
   |||||
DB 315 AGACTGACCCAGGCGCAGCACTAAAGAAAGAAATTAATGCAAGTGGTCTTAGCTCCACTTTA 374

QY 361 GCTTTAATAATATGTTTCATATATATCTCTGCTTTTCTCTATATAAAGAGCTTGT 420
   |||||
DB 375 GCTTTAATAATATGTTTCATATATATCTCTGCTTTTCTCTATATAAAGAGCTTGT 434

QY 421 ATTTTCATTTGAAGCAGAGCGGCAACACACACAGAACCTCCCTGCTTACAAACC 476
   |||||
DB 435 ATTTTCATTTGAAGCAGAGCGGCAACACACACAGAACCTCCCTGCTTACAAACC 490

RESULT 3
ABX09816
ID ABX09816 standard; DNA; 1057 BP.
XX
AC ABX09816;
XX
DT 22-JAN-2003 (first entry)
XX
DE Promoter sequence #1 useful in method for nematode control.
XX
KW RNAi molecule; double-stranded interfering RNA; nematode control;
KW RNA mediated interference; mRNA transcript; nematode gene; growth;
KW development; parasitism; reproduction; RNAi vector; mRNA translation;
KW nematode inhibitor; agricultural industry; anti-nematode; promoter; ds.
XX
OS Unidentified.
XX
PN WO200196584-A2.
XX
```





Query Match 100.0%; Score 476; DB 6; Length 1071;  
Best Local Similarity 100.0%; Pred. No. 2.4e-123;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTATGCTGATGACAGATTAATCTTAATAGCAATATTTCAGAAATTAATCAAGAGAAA 60  
DB 563 AACTATGCTGATGACAGATTAATCTTAATAGCAATATTTCAGAAATTAATCAAGAGAAA 622  
QY 61 GAATTAATAACTCTTTTCAAGATATGAGCCCGCTTTTCAAGTGGCCAGCTAGCTATCACT 120  
DB 623 GAATTAATAACTCTTTTCAAGATATGAGCCCGCTTTTCAAGTGGCCAGCTAGCTATCACT 682  
QY 121 GAAAGACAGCAAGCAATGTGTCTGATGACACAGAACCAATCTTTTGACGAGATGT 180  
DB 683 GAAAGACAGCAAGCAATGTGTCTGATGACACAGAACCAATCTTTTGACGAGATGT 742  
QY 181 GAAGCAGCCAGAGTGGTCCACAGAGCCGACCTCAGAAAGGACATCTTACCGACACAGAA 240  
DB 743 GAAGCAGCCAGAGTGGTCCACAGAGCCGACCTCAGAAAGGACATCTTACCGACACAGAA 802  
QY 241 AAAGCAACCCAGCTCATCTCAACATGTAGACTGTCTGTTATGCGTGGCTGAAGATA 300  
DB 803 AAAGCAACCCAGCTCATCTCAACATGTAGACTGTCTGTTATGCGTGGCTGAAGATA 862  
QY 301 AGACTGACCCAGCCAGCCAGCTCAAGAGAAATTAATGCAAGTGGTCTAGCTCCACTTTA 360  
DB 863 AGACTGACCCAGCCAGCCAGCTCAAGAGAAATTAATGCAAGTGGTCTAGCTCCACTTTA 922  
QY 361 GCTTTAATAATATGTTTTCATTAATTAATTTCTGCTTTTCTCTCTATATAAGAGCTTGT 420  
DB 923 GCTTTAATAATATGTTTTCATTAATTAATTTCTGCTTTTCTCTCTATATAAGAGCTTGT 982  
QY 421 ATTTCATTTGAGCAGAGCGGAAACACACACAGAAACCTCCCTGCTTACAAACC 476  
DB 983 ATTTCATTTGAGCAGAGCGGAAACACACACAGAAACCTCCCTGCTTACAAACC 1038

RESULT 5  
AAD39134  
ID AAD39134 standard; DNA; 651 BP.  
XX AC AAD39134;  
XX DT 04-OCT-2002 (first entry)  
XX DE Strawberry vein banding virus (SVBV) promoter DNA, sesVBV.  
XX DW Strawberry vein banding virus promoter; transgenic plant; transgenic; ds.  
XX OS Strawberry vein banding virus.  
XX PN WO200240691-A2.  
XX PD 23-MAY-2002.  
XX PF 30-OCT-2001; 2001WO-US047964.  
XX PR 01-NOV-2000; 2000US-0245354P.  
XX XX (MAXY-) MAXYGEN INC.  
XX PI Wu G, McBride K;  
XX DR WPI; 2002-479908/51.  
XX PT Novel isolated or recombinant nucleic acid comprising strawberry vein  
XX PT banding virus promoter operably linked to a heterologous polynucleotide,  
XX PT useful for expressing the heterologous polynucleotide in a plant cell.  
XX PS Claim 1; Page 47-48; 48pp; English.  
XX XX The invention relates to an isolated/recombinant nucleic acid comprising  
XX CC a strawberry vein banding virus (SVBV) promoter operably linked to a

CC heterologous polynucleotide. The invention is useful for expressing a  
CC heterologous polynucleotide in a plant cell, by introducing the nucleic  
CC acid into a plant cell, where the plant cell is present within a plant  
CC and Agrobacterium is used to introduce the isolated nucleic acid into the  
CC cell. The invention is useful to drive gene expression in plant cells and  
CC transgenic plants, for transforming plant cells and producing transgenic  
CC plants, for compensating an endogenous protein at higher than normal levels,  
CC plant, for expressing a novel gene in a plant, and to suppress expression of  
CC endogenous plant genes. The present sequence is sesVBV promoter DNA  
XX  
SQ Sequence 651 BP; 213 A; 167 C; 132 G; 139 T; 0 U; 0 Other;

Query Match 77.5%; Score 369; DB 6; Length 651;  
Best Local Similarity 100.0%; Pred. No. 2.2e-93;  
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 GCTAGCTATCACTGAAAAGACAGCAGCAATGGTGTCTCGATGCACAGAACCACTCT 167  
DB 278 GCTAGCTATCACTGAAAAGACAGCAGCAATGGTGTCTCGATGCACAGAACCACTCT 337  
QY 168 TTGCAGCAGATGTGAAGCAGCCAGAGTGTCTCACAAGACGCACTCAGAAAAGCATCTTC 227  
DB 338 TTGCAGCAGATGTGAAGCAGCCAGAGTGTCTCACAAGACGCACTCAGAAAAGCATCTTC 397  
QY 228 TACCGACAGAAAAGACCAACAGCTCATCTCAACATGTAGACTGTGTTATGG 287  
DB 398 TACCGACAGAAAAGACCAACAGCTCATCTCAACATGTAGACTGTGTTATGG 457  
QY 288 TCGGCTGAAGATAAGACTGACCCAGCCAGCAGCACTTAAAGAGAAAATAATGCAAGTGGTCC 347  
DB 458 TCGGCTGAAGATAAGACTGACCCAGCCAGCAGCACTTAAAGAGAAAATAATGCAAGTGGTCC 517  
QY 348 TAGCTCCACTTTAGCTTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 407  
DB 518 TAGCTCCACTTTAGCTTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 577  
QY 408 ATAAAGAGCTTCTATTTTTCATTTGAAGCAGAGCGGAAACACACACAGAAACCTCCCTGC 467  
DB 578 ATAAAGAGCTTCTATTTTTCATTTGAAGCAGAGCGGAAACACACACAGAAACCTCCCTGC 637  
QY 468 TTACAAACC 476  
DB 638 TTACAAACC 646

RESULT 6  
AAD39135  
ID AAD39135 standard; DNA; 701 BP.  
XX AC AAD39135;  
XX DT 04-OCT-2002 (first entry)  
XX DE Strawberry vein banding virus (SVBV) promoter DNA, lesVBV.  
XX DW Strawberry vein banding virus promoter; transgenic plant; transgenic; ds.  
XX OS Strawberry vein banding virus.  
XX PN WO200240691-A2.  
XX PD 23-MAY-2002.  
XX PF 30-OCT-2001; 2001WO-US047964.  
XX PR 01-NOV-2000; 2000US-0245354P.  
XX XX (MAXY-) MAXYGEN INC.  
XX PI Wu G, McBride K;  
XX DR WPI; 2002-479908/51.

XX	Novel isolated or recombinant nucleic acid comprising strawberry vein
PT	banding virus promoter operably linked to a heterologous polynucleotide,
PT	useful for expressing the heterologous polynucleotide in a plant cell.
XX	Claim 1; Page 48; 48pp; English.
XX	The invention relates to an isolated/recombinant nucleic acid comprising
CC	a strawberry vein banding virus (SVBV) promoter operably linked to a
CC	heterologous polynucleotide. The invention is useful for expressing a
CC	heterologous polynucleotide in a plant cell, by introducing the nucleic
CC	acid into a plant cell, where the plant cell is present within a plant
CC	and Agrobacterium is used to introduce the isolated nucleic acid into the
CC	cell. The invention is useful to drive gene expression in plant cells and
CC	transgenic plants, for transforming plant cells and producing transgenic
CC	plants, for compensating for missing or altered gene expression in a
CC	plant, for expressing an endogenous protein at higher than normal levels,
CC	for expressing a novel gene in a plant, and to suppress expression of
CC	endogenous plant genes. The present sequence is lesVBV promoter DNA
XX	
SQ	Sequence 701 BP; 222 A; 176 C; 139 G; 164 T; 0 U; 0 Other;
	Query Match 77.5%; Score 369; DB 6; Length 701;
	Best Local Similarity 100.0%; Pred. No. 2.2e-93;
	Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	108 GCTAGCTATCACTGAAAAGACAGCAAGACAATGGTGCTCGATGCACCAAGAACCCACATCT 167
DB	319 GCTAGCTATCACTGAAAAGACAGCAAGACAATGGTGCTCGATGCACCAAGAACCCACATCT 378
QY	168 TTTCAGCAGATGTGAAGCAGCAGAGTGGTCCACAAGCGCACTCAGAAAAGGCAATCTTC 227
DB	379 TTTCAGCAGATGTGAAGCAGCAGAGTGGTCCACAAGCGCACTCAGAAAAGGCAATCTTC 438
QY	228 TACCGACACAGAAAAGACACACCACAGCTCATCTCCAAACATGTAGACTGTGTTATGG 287
DB	439 TACCGACACAGAAAAGACACACCACAGCTCATCTCCAAACATGTAGACTGTGTTATGG 498
QY	288 TCGGCTGAAGATAAGACTGACCCAGGCGCAGCACTAAAGAAAGAAATAATGCAAGTGGTCC 347
DB	499 TCGGCTGAAGATAAGACTGACCCAGGCGCAGCACTAAAGAAAGAAATAATGCAAGTGGTCC 558
QY	348 TAGCTCCACTTTAGCTTTTAATTAATTATGTTTCTATTATTTCTGCTTTTGCTCTCTAT 407
DB	559 TAGCTCCACTTTAGCTTTTAATTAATTATGTTTCTATTATTTCTGCTTTTGCTCTCTAT 618
QY	408 ATAAAGAGCTTGTAATTTTCATTGAGCGCAGAGCGCAACACACACAGAACCTCCCTGC 467
DB	619 ATAAAGAGCTTGTAATTTTCATTGAGCGCAGAGCGCAACACACACAGAACCTCCCTGC 678
QY	468 TTACAAACC 476
DB	679 TTACAAACC 687
RESULT 7	
ABL26902/c	
ID	ABL26902 standard; DNA; 3294 BP.
XX	
AC	ABL26902;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster genomic polynucleotide SEQ ID NO 32179.
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide;
XX	pharmaceutical; gene; ds.
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.

[illegible]

PA (CHIR ) CHIRON CORP.  
 XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC, Scarlato V;  
 PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;  
 PI Rappuoli R, Pizzo M;  
 XX WPI; 2000-318079/27.  
 DR Isolated nucleotide sequences of *Neisseria meningitidis* which can be used  
 PT in the diagnosis and treatment of *N. meningitidis* infection and other  
 PT Neisserial infections, for example, *N. gonorrhoea*.  
 XX

PS Claim 7; Page 330-353; 1760pp; English.

XX The present invention describes methods of obtaining immunogenic proteins  
 CC from *Neisseria* genomic sequences. AAA81453 to AAA82414 represent  
 CC specifically claimed *Neisseria meningitidis* genomic DNA sequences;  
 CC AAA81260 to AAA81303 and AAA82562 to AAA82563 represent *Neisseria* DNA  
 CC sequences and their corresponding proteins; AAA81254 to AAA81259 and  
 CC AAA81304 to AAA81321 represent PCR primers used in the isolation of  
 CC *Neisseria meningitidis* DNA sequences; and AAA81322 to AAA81452 represent  
 CC *Neisseria meningitidis* MenB polynucleotide ORF sequences, which are all  
 CC used in the exemplification of the present invention. The nucleic acid  
 CC sequences, protein sequences, and antibodies against them, can be used in  
 CC the manufacture of a composition. The composition can be used as a  
 CC medicament (or in the manufacture of a medicament) for treating,  
 CC preventing or diagnosing infection due to *Neisseria* bacteria. For  
 CC example, some of the identified proteins could be components of vaccines  
 CC against *Meningococcus* B; against all serotypes; and/or against all  
 CC pathogenic *Neisseria*. Identification of sequences from the bacterium  
 CC will also facilitate production of biological probes, particularly  
 CC organism-specific probes. Attempts to make efficacious *Meningococcus* B  
 CC vaccines have failed mainly due to antigen tolerance. Multivalent  
 CC vaccines have also been tried but none have successfully overcome  
 CC antigenic variability. The provision of further, complete sequences may  
 CC provide an opportunity to identify secreted or surface exposed proteins  
 CC that may be presumed targets for the immune system and which are not  
 CC antigenically variable or at least more conserved than other more  
 CC variable regions  
 XX

SQ Sequence 78845 BP; 19372 A; 19372 C; 19042 G; 19318 T; 0 U; 2 Other;

Query Match 7.9%; Score 37.6; DB 3; Length 78845;  
 Best Local Similarity 48.2%; Pred. No. 14;  
 Matches 106; Conservative 0; Mismatches 114; Indels 0; Gaps 0;  
 QY 217 AAGGATCTTCCGACACAGAAACACACACAGCTCATCATCAACATGTAGACT 276  
 DB 11186 AAGGATCTTCCGAAATTTGAAAGAAATAGAGCCGAGGATTTGATGGATGATTTACC 11245  
 QY 277 GTCGTTATGCGCTGAACATAGACTGACCCGAGCCGAGCTAAAGAAAGAAATAT 336  
 DB 11246 GTGCTTTGATCGGTGGAAATTTCTGAATTCGCTATAGTAAAGAAAGACAAAAG 11305  
 QY 337 GCAAGTGTCTAGCTCCACTTTAGCTTTAATATATGTTTCATTTATTTCTCTGCTT 396  
 DB 11306 GGTTTTGGTAAGTACCAAAATCAAAACAGTTGATTTGATTCGGCAGATATGTTCTTG 11365  
 QY 397 TTGCTCTCTATATAAGAGCTTGTATTTTCATTTGAAGGC 436  
 DB 11366 ATGATCTTTTCTGAAATGCTTACAGCTTCATATAGGC 11405

RESULT 9  
 AAA81490 08/C  
 Continuation (9 of 15) of AAA81490 from base 800001 (N. meningitidis B full length genome)  
 WP Sequence split into 15 fragments LOCUS AAA81490 Accession AAA81490  
 WP Fragment Name Begin End  
 WP AAA81490\_00 1 110000  
 WP AAA81490\_01 100001 210000  
 WP AAA81490\_02 200001 310000  
 WP AAA81490\_03 300001 410000  
 WP AAA81490\_04 400001 510000

WP AAA81490\_05 500001 610000  
 WP AAA81490\_06 600001 710000  
 WP AAA81490\_07 700001 810000  
 WP AAA81490\_08 800001 910000  
 WP AAA81490\_09 900001 1010000  
 WP AAA81490\_10 1000001 1110000  
 WP AAA81490\_11 1100001 1210000  
 WP AAA81490\_12 1200001 1310000  
 WP AAA81490\_13 1300001 1410000  
 WP AAA81490\_14 1400001 1437668  
 Query Match 7.9%; Score 37.6; DB 3; Length 110000;  
 Best Local Similarity 48.2%; Pred. No. 16;  
 Matches 106; Conservative 0; Mismatches 114; Indels 0; Gaps 0;  
 QY 217 AAGGATCTTCTACGACACAGAAACACACAGCTCATCATCAACATGTAGACT 276  
 DB 83383 AAGGATCTCGGAATTTGAAAGAAATAGAGCCGAGGATTTGATGGATGATTTACC 83324  
 QY 277 GTCGTTATGCGCTGAAGATAAGACTGACCCGAGCCGAGCTAAAGAAAGAAATAT 336  
 DB 83323 GTGCTTTGATTCGGTTGGAAATTTCTGAATTCGCTATAGTAAAGAAAGACAAAAG 83264  
 QY 337 GCAAGTGTCTAGCTCCACTTTAGCTTTAATATATGTTTCATTTATTTCTCTGCTT 396  
 DB 83263 GGTTTTGGGTAAGTACCAAAATCAAAACAGTTGATTTGATTCGGCAGATATGTTCTTG 83204  
 QY 397 TTGCTCTCTATATAAGAGCTTGTATTTTCATTTGAAGGC 436  
 DB 83203 ATGATCTTTTCTGAAATGCTTACAGCTTCATATAGGC 83164

RESULT 10

AAF21608/C  
 ID AAF21608 standard; DNA; 349980 BP.

XX AAF21608;  
 AC AAF21608;  
 XX DT 13-MAR-2001 (first entry)  
 XX DE *Neisseria meningitidis* B nucleotide sequence SEQ ID NO:109.

XX *Neisseria meningitidis*; *Neisseria gonorrhoeae*; immunogenic; vaccine;  
 KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;  
 KW ds.  
 XX *Neisseria meningitidis*.  
 OS WO2000066791-A1.  
 PN 09-NOV-2000.  
 PD 08-MAR-2000; 2000WO-US005928.  
 PF 30-APR-1999; 99US-0132068P.  
 PR 08-OCT-1999; 99WO-US023573.  
 PR 26-FEB-2000; 2000GB-00004695.  
 XX (CHIR ) CHIRON CORP.  
 PA (GENO-) INST GENOMIC RES.

XX Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC;  
 PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;  
 PI Rappuoli R, Frazer CM, Grandi G;  
 XX WPI; 2000-647603/62.  
 DR *Neisseria meningitidis* B full length genome sequence and open reading  
 PT frames are used to detect, treat and prevent *Neisseria* infections.  
 XX Claim 7; Appendix A; 692pp; English.  
 XX The present invention describes the full length genome of *Neisseria*

CC meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613  
CC represent fragments of the NMB genomic sequence, as the sequences was too  
CC long to go in a record on its own it was split into 8 sequences which  
CC overlap each other at the beginning and end of each sequence by 49980 bp  
CC (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of  
CC AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of  
CC AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins  
CC given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR  
CC primers which are used in the exemplification of the present invention.  
CC The NMB genome and fragments from it have antibacterial activity, and can  
CC be used in vaccines and gene therapy. Neisseria nucleic acids, proteins  
CC and/or antibodies which binds to the proteins can be used in compositions  
CC for treating or preventing infection due to Neisseria bacteria or as a  
CC diagnostic reagent for detecting the presence of Neisseria bacteria  
CC of antibodies raised to Neisseria bacteria. Computers, computer memory,  
CC computer storage medium or computer databases can be used in a search to  
CC identify open reading frames (ORFs) or coding sequences within the NMB  
CC genome. The DNA sequences provide further opportunities to find antigenic  
CC or immunogenic proteins which are more effective in vaccines than the  
CC outer membrane proteins currently used  
XX  
SQ Sequence 349980 BP; 82523 A; 82940 C; 96712 G; 87805 T; 0 U; 0 Other;  
Query Match 7.9%; Score 37.6; DB 3; Length 349980;  
Best Local Similarity 48.2%; Pred. No. 24;  
Matches 106; Conservative 0; Mismatches 114; Indels 0; Gaps 0;  
QY 217 AAGGCATCTTCACGACACAGAAAGACACACAGCTCATCTCAACATGTAGACT 276  
Db 283384 AAGGATGCGGAATTGGAAAGAAATAGAGCCGCGAGATTGTGATGATGTTTACC 283325  
QY 277 GTGCTTATCGTGGCTGAAGATAAGACTGACCCAGCCGACACTAAAGAAAGAAATAT 336  
Db 283324 GTGCTTATCGTGGCTGAAGATAATCTGAAATATTCGTATAGTAAAGAGCAAAAG 283265  
QY 337 GCAAGTGGTCTAGCTCCACTTAGCTTTAAATATATCTTTCATTTATTTCTGCTT 396  
Db 283264 GGTTTTGGTGAAGTACCAAAATCAAAACAGTTGTATTTGATCGGCAATATGTTCTG 283205  
QY 397 TTGCTCTCTATATAAAGAGCTTGTATTTTCATTTGAAGGC 436  
Db 283204 ATGATCTTTTCTGAATGCTTACAGCTTCATATAGGC 283165  
RESULT 11  
AAS46724/c  
ID AAS46724 standard; DNA; 6012 BP.  
XX  
AC AAS46724;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Tumour suppressor gene derived chemically modified sequence #447.  
XX  
KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;  
KW tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;  
KW cytosine methylation; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200168912-A2.  
XX  
FD 20-SEP-2001.  
XX  
FF 15-MAR-2001; 2001WO-EP002955.  
XX  
PR 15-MAR-2000; 2000DE-01013847.  
PR 06-APR-2000; 2000DE-01019058.  
PR 07-APR-2000; 2000DE-01019173.  
PR 30-JUN-2000; 2000DE-01032529.  
PR 01-SEP-2000; 2000DE-01043826.  
XX  
PA (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;  
XX WPI; 2001-602752/68.  
XX  
XX Fragments of chemically modified genes associated with tumor suppressor  
XX genes and oncogenes, useful in designing primers and probes for analyzing  
XX diseases associated with cytosine methylation state e.g. cancer.  
XX  
XX Claim 1; SEQ ID NO 447; 27pp; English.  
XX  
XX The invention relates to a nucleic acid comprising a sequence of 18  
XX bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with  
XX bisulphite, of genes associated with tumour suppression and oncogenes  
XX having a sequence taken from 536 (actually 533 since numbers 408, 458 and  
XX 500 are missing from the sequence listing) sequences (Ss) and sequences  
XX complementary to (Ss). The nucleic acid may be a peptide nucleic acid-  
XX oligomer (PNA) of at least 9 nucleotides and may form part of a set of  
XX probes for detecting the cytosine methylation state and/or single  
XX nucleotide polymorphisms and also to be used in an array for analysing  
XX diseases associated with CpG dinucleotides e.g. cancers and tumours. The  
XX probes can also be used in a method for ascertaining genetic and/or  
XX epigenetic parameters for the diagnosis and/or therapy of existing  
XX diseases or the predisposition to specific diseases, by analysing  
XX cytosine methylations. The parameters may be compared to another set of  
XX genetic and/or epigenetic parameters, the differences serving as basis  
XX for diagnosis and/or prognosis events which are disadvantageous to  
XX patients. The present sequence is one of the 533 genomic sequences  
XX derived from tumour suppressor genes and oncogenes. Note: The sequence  
XX data for this patent did not form part of the printed specification, but  
XX was obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
SQ Sequence 6012 BP; 1335 A; 107 C; 1758 G; 2812 T; 0 U; 0 Other;  
Query Match 7.7%; Score 36.8; DB 4; Length 6012;  
Best Local Similarity 47.4%; Pred. No. 8.7;  
Matches 110; Conservative 0; Mismatches 122; Indels 0; Gaps 0;  
QY 34 AATTATTCAGATTAATCAAGGAGAGAAATTAATTAATCTCTTCAGATATGAAGCCGC 93  
Db 861 ACTCATTTTACAAAAAACAACAAACAAACAACTAAATAACTTACCACAAATACAAACACC 802  
QY 94 TTTCAAGTGGCCAGCTAGCTATCACTGAAAAGACAGCAAGCAATGTGTCTCGATCA 153  
Db 801 AAACCCACTACCAACAAACCTATACCCAAATACAAAAAATAATATCTCAAAAA 742  
QY 154 CCAGAACCAATCTTTTGCGCAGATGTGACGACGAGGTGGTCCACAGACCATCA 213  
Db 741 CAAAAAATATAACGACAAAAAACAACAAAAAACCCTACATAAAAAAACCACCA 682  
QY 214 GAAGAAGCATCTTCTACCGACACAGAAAAAGACCAACCACTCATCATCA 265  
Db 681 ATAAAAACACACACACACACACCCCAATANACACACACACACCCCAATAAACA 630  
RESULT 12  
ABL32655/c  
ID ABL32655 standard; DNA; 17211 BP.  
XX  
AC ABL32655;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Human immune system associated gene SEQ ID NO: 628.  
XX  
KW Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antiandemic; cytostatic; neotropic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmologic;  
KW antirheumatic; antiarthritis; antidiabetic; antipsoriatic;  
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;

```

KW ds.
XX Homo sapiens.
XX WO200200928-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP007537.
XX
XX 30-JUN-2000; 2000DE-01032529.
XX
XX 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful for
XX diagnosis and treatment of diseases associated with abnormal cytosine
XX methylation.
XX
XX Claim 1; SEQ ID NO 628; 32bp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX diseases. The present sequence is a gene of the invention
XX
XX Sequence 17211 BP; 5518 A; 180 C; 3351 G; 8162 T; 0 U; 0 Other;
XX
XX Query Match 7.7%; Score 36.6; DB 6; Length 17211;
XX Best Local Similarity 56.1%; Pred. No. 15;
XX Matches 69; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
XX
XX QY 17 AGATAATCTTAATAGCAATTAATTCAGAAATTAATCAAGGAGAAAGAAATTAATTAATCTTT 76
XX
XX Db 7211 AAATAATCTCTCCACCTCATCTTCCGAATAACTAAATAATTAATAATTTTCAAAAACCTT 7152
XX
XX QY 77 CAGAAATATGAAGCCCGTTTACAGTGCAGCTAGCTATCATCTGAAAGACAGCAGAC 136
XX
XX Db 7151 TATATAATATATCTCTATATAATTTCCCAACTCTCTATAACTATAAATTTCTTATAA 7092
XX
XX QY 137 AAT 139
XX
XX Db 7091 AAT 7089
XX
XX RESULT 13
XX ABZ38820
XX ID ABZ38820 standard; DNA; 285 BP.
XX
XX AC ABZ38820;
XX
XX 07-MAR-2003 (first entry)
XX
XX N. gonorrhoeae nucleotide sequence SEQ ID 2229.
XX
XX Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX
XX Neisseria gonorrhoeae.
XX
XX WO200279243-A2.
XX
XX 10-OCT-2002.
XX
XX 12-FEB-2002; 2002WO-1B002069.
XX
XX
XX
```

```

PR 12-FEB-2001; 2001GB-00003424.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Fontana MR, Pizza M, Masignani V, Monaci E;
XX
XX WPI; 2003-058415/05.
XX
XX P-PSDB; ABP77850.
XX
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX medicament for treating or preventing N. gonorrhoeae infection.
XX
XX Disclosure; Page 349; 815pp; English.
XX
XX The present invention relates to proteins from Neisseria gonorrhoeae.
XX Also disclosed are the nucleic acid molecules encoding the proteins and
XX antibodies that specifically bind to the proteins. The composition
XX comprising the protein, nucleic acid or antibody is useful for the
XX manufacture of a medicament for treating or preventing N. gonorrhoeae
XX infection, this may be in the form of a vaccine or gene therapy.
XX Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
XX molecules of the invention
XX
XX Sequence 285 BP; 91 A; 40 C; 64 G; 90 T; 0 U; 0 Other;
XX
XX Query Match 7.6%; Score 36.4; DB 7; Length 285;
XX Best Local Similarity 47.7%; Pred. No. 3.5;
XX Matches 106; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
XX
XX QY 235 ACAGAAAAGACAACACAGCTCATCTCCACATGATGATAGCTGCTGTTATCGTCGGCTG 294
XX
XX Db 5 AAAAGAAAATAGAGCGCGAGGATTGATGGATGATTTTACCGTCTTTGATTGGTTG 64
XX
XX QY 295 AAGATAAGACTGACCCCGAGCGAGCACTAAAGAGAAATAATGCAAGTGGTCTAGCTCC 354
XX
XX Db 65 GAAATATTCTGAATTTCCGTATAGTAGAAGAGCAAAAGGGTTTGGTAAATACCA 124
XX
XX QY 355 ACTTAGCTTTAAATTAATTAATGTTTCATTAATTAATTCCTGCTCTCTATATAAGA 414
XX
XX Db 125 AAATCAAAACAGTTGTAATTTGATTGGCAGATATGCTCTCTGATGATCTTTTCTGAAAT 184
XX
XX QY 415 GCTTCTATTTTCATTTGAAGCGCAGCGGCAACACACACACAG 456
XX
XX Db 185 GCTTACAGTCTTCATATAAGGCTTATTTCTGAACTGAACCTG 226
XX
XX RESULT 14
XX AAD53224_2
XX Continuation (3 of 6) of AAD53224 from base 200001 (Human chromosome 3 q-arm breakpoint :
XX WP Sequence split into 6 fragments LOCUS AAD53224 Accession AAD53224
XX WP Fragment Name Begin End
XX WP AAD53224_0 1 110000
XX WP AAD53224_1 100001 210000
XX WP AAD53224_2 200001 310000
XX WP AAD53224_3 300001 410000
XX WP AAD53224_4 400001 510000
XX WP AAD53224_5 500001 567571
XX
XX Query Match 7.6%; Score 36.2; DB 7; Length 110000;
XX Best Local Similarity 56.2%; Pred. No. 38;
XX Matches 68; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
XX
XX QY 317 AGCAATAAGAGAAATAATGCAAGTGGTCTAGCTCCACTTAGCTTAATTAATATGT 376
XX
XX Db 103031 AGCACAGGCATTTTAAAGTAACTGCTCTGCTCATTTGACATATAATTTATTGA 103090
XX
XX QY 377 TCAATTAATATCTCTGCTTTTGTCTCTATATAAGAGCTTGTATTTCATTGAAGGC 436
XX
XX Db 103091 GTCAATCTTATTGACTTCTTTTGAACCTCTATCCACAAAACCTGAGGCGAATAGAGG 103150
XX
XX QY 437 A 437
XX
XX Db 103151 A 103151
XX
```



